Determining the mechanisms driving the invasion success of the smallmouth bass (*Micropterus dolomieu*, Lacepède 1802)

by

Genevieve Diedericks

Dissertation presented for the degree of
Doctor of Philosophy in the
Faculty of Science at
Stellenbosch University

Supervisor: Professor C. Hui

Co-supervisors: Professor S. von der Heyden
Doctor O.L.F. Weyl

December 2017

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2017

ABSTRACT

Biological invasions represent a significant threat to the Earth's biota, including freshwater ecosystems. Whilst various studies have been conducted on invasive species in the hopes of establishing a set of general guidelines to facilitate our understanding of the invasion pathway and the features (organismal or environmental) that may assist in a species' invasion success, predictions and generalisations remain challenging and limited. While general guidelines and unified frameworks are essential in furthering our understanding of biological invasions, exploring the genetic and phenotypic variation in a species and the influence of extrinsic factors during adaptation to novel environments could provide insight into a species' invasion success. By utilising an array of methods, spanning different disciplines, this dissertation evaluates the potential mechanisms driving a species' invasive success using the smallmouth bass, *Micropterus dolomieu,* as model organism. First, I attempt to unravel the role of genetic diversity in a species' invasive success, by assessing the genetic differentiation and diversity within *M. dolomieu* populations in the invaded South African (SA) range, and examine how the genetic diversity may change over time in both native (USA) and invasive ranges (Chapter 2). By conducting a three-way comparison on two mitochondrial (mtDNA) and nine microsatellite loci for a total of 572 *M. dolomieu* specimens, representing the contemporary invasive SA range, contemporary native USA range and the historic native USA range (dating back to the period of introduction into SA), I reveal the presence of elevated levels of genetic diversity for the invasive SA range. The levels of genetic diversity for both the contemporary native and invasive ranges are, however, substantially lower than those of the historic native range, suggesting that both contemporary populations experienced a recent genetic bottleneck followed by a rapid population expansion.

As admixture, and more specifically introgressive hybridisation, may contribute to these elevated levels of genetic diversity observed within the invaded SA range, I subsequently test the hypothesis that hybridisation and introgression can occur between two invasive species in a novel invaded range (Chapter 3). Using two notorious freshwater invaders (*M. dolomieu* and *M. salmoides*), I assess the level of introgressive hybridisation between the two species, using two mtDNA and nine microsatellite loci. Despite large uncorrected pairwise distances being observed between the two species, unidirectional mitochondrial introgression was detected, suggesting that

introgressive hybridisation may play a pivotal role in the successful establishment and spread of alien invasive species upon introduction.

The remaining chapters focus on the role of morphological variation in response to environmental variation in the novel invaded range. First, I test the hypothesis that environmental variation drives morphological changes in phenotype (Chapter 4). Analyses of linear and geometric morphometrics, as well as environmental variables, show a strong correlation between body depth and flow regime, with streamlined fish inhabiting high-flow environments. In addition, the presence of variation in gape size among localities suggests a link between cranial morphology and prey composition associated with substrate type. These results support the idea that similar environments have convergent phenotypes and highlight the importance of phenotypic plasticity in facilitating the successful colonisation, establishment and spread of invasive species.

However, as gene flow may erase phenotypic variation associated with plasticity, I continue by exploring the genetic basis to local adaptation (Chapter 5). By combining linear morphometrics and genotypic data, I demonstrate slight population structuring among sampled localities, corresponding to three distinct sections of the river (i.e., tributary, impoundment and mainstem). However, the presence of high levels of gene flow observed among populations suggesting that both local adaptation and phenotypic plasticity may play a key role.

In conclusion, *M. dolomieu* appears to be a successful invader that fully exploits genetic and phenotypic variability to invade, persist and establish in an array of non-native environments.

ACKNOWLEDGEMENTS

Numerous people contributed to the successful completion of this thesis and I would like to acknowledge them here.

Firstly, I would like to thank my supervisors, Cang Hui, Sophie von der Heyden and Olaf Weyl for their constant support and guidance throughout my thesis. I would like to extend a special word of thanks to Cang for seeing potential in my proposed project and giving me ample freedom to grow, discover and drive my own PhD. Not many academics would climb onboard a project envisioned by a student and support them with insightful comments and ideas, guidance and financial assistance. Cang, you gave me the freedom to dream big (something zoologists and geneticists can't usually do due to funding limitations), not only with regards to my thesis but also in allowing me to guide an honours student during my PhD. The knowledge and experience I gained can never be erased – thank you.

A special thanks to my fiancé Chris Broeckhoven and my two fur-children Cleo(patra) and Xena. The three of you made it worthwhile getting up at the end! Chris, thank you for all the love, support, encouragement, brainstorming, statistic help, proof-reading, questioning some of my wild ideas and endless field assistance! Without your insatiable love for fishing, even if it means getting cut by *Palmiet* (*Prionium serratum*), and the dogs' love for swimming, kayaking and the fish itself, we might still be busy collecting bass!

I am grateful to my family for their continued support throughout my PhD. The family trips collecting bass from dusk till dawn, wading through the Kouga river, catching fry with a make-shift net and all the arguments about who-cast-over-who's line or who-stole-who's fish are memories I'll always cherish. To my almost- family in Belgium, thank you for your support throughout. Having my (almost) sister-in-law finish her PhD thesis just a few months before me really helped me stay motivated toward the end.

Starting this PhD, I had never worked with microsatellites before and so needed quite some guidance. Thank you Romina Henriques and Sara Andreotti from the Evolutionary Genomics

Group (EGG lab), Department of Botany and Zoology, Stellenbosch University, for your invaluable assistance in this regard. Without your knowledge, assistance and paper recommendations I would probably still be trying to score them! Thanks to the EGG lab (past and present members) for being my family for the past seven years, particularly Nina, Romina, Sara, Louisa, Lisa and Ethel. The numerous philosophical chats, life guidance and support from so many of you made many difficult times bearable and helped me maintain my sanity throughout this time. Cleo and Xena will of course miss all the cuddles and attention from all of you too.

Thanks to the following people and institutions for their sample contributions: Erling Holm (Royal Ontario Museum), Douglas Nelson (University of Michigan Museum of Zoology), Jeff Williams (National Museum of Natural History), Mark Sabaj Pérez (Academy of Natural Sciences), Mark Kibbey (Ohio State University Museum), South African Institute for Aquatic Biodiversity (SAIAB), Jason Barnucz (Fisheries and Oceans Canada), Wil Wegman (Ontario Ministry of Natural Resources and Forestry, Aurora District), Gene Gilliland (B.A.S.S conservation director), Rich Carter (Ohio Department of Natural Resources), Randy Jackson (Department of Natural Resources, Cornell University), Jeff Loukmas (New York State Department of Environmental Conservation), Jacques Aproskie and Craig Fraser (both professional South African bass anglers). A special word of thanks to Dawie Burger from Driehoek farm, Cederberg, for collecting a substantial portion of the largemouth bass used in Chapter 3. Last but not least, thank you to Francois Roux from Mpumalanga Parks Board for all your assistance in obtaining sampling permits, arranging the eager anglers, and coordinating all the sampling trips to the Blyde River Canyon, a world heritage site. Without your help, I would not have been able to collect specimens in this magical place.

Furthermore, I'd like to acknowledge Dean Impson, Martine Jordaan, Bruce Paxton and Riaan van der Walt for advice and insight into various bass related topics (sampling methods and equipment, river systems etc.) and Richard Peel and Geraldine Taylor from SAIAB for preparing the otoliths for scoring. Thank you too to the Central Analytical Facilities (CAF), Stellenbosch University, for genotyping and sequencing all specimens used in this study. The support staff of the Centre for Invasion Biology, Department of Botany & Zoology and Mathematical Sciences (Stellenbosch University), especially Mathilda van der Vyver, Janine Basson, Mari Sauerman and

Vanessa du Plessis, for assisting with various requests and problems. The Research Ethics Committee of the Faculty of Science (Stellenbosch University) for ethical clearance and the Western Cape, Eastern Cape and Mpumalanga provincial conservation authorities for collection permits.

I would like to thank the Centre of Excellence for Invasion Biology (C.I.B) and the Department of Botany and Zoology for hosting me. Lastly, I am grateful to the DST-NRF Centre of Excellence for Invasion Biology for a PhD scholarship and funding from the National Research Foundation (NRF), awarded to Cang Hui and Olaf Weyl, for financial support during my study.

Finally, I would like to thank the three anonymous reviewers for their insightful comments and suggestions. Their contributions greatly improved the final version of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

GENERAL INTRODUCTION

Biological invasion, a process whereby the Earth's biota is being homogenised (Elton 1958), is a leading threat to global biodiversity (Lodge 1993; Vander Zanden et al. 2004). Induced by human-mediated dispersal and global environmental change (Moyle & Light 1996; Heger & Trepl 2003; Wilson et al. 2009; Ricciardi & MacIsaac 2011; Ferrari et al. 2014), biological invasions represent a significant threat, particularly to freshwater ecosystems (Dudgeon et al. 2006; Ricciardi 2007; Blanchet 2012; Arismedi et al. 2014), though in most instances, successful invaders were intentionally introduced due to their perceived benefit to society (García- Berthou 2007; Leprieur et al. 2009; Marr et al. 2010; Ehrenfeld 2010; Arismedi et al. 2014). A plethora of studies covering various research fields have been conducted on a variety of invasive species in the hopes of establishing a set of general guidelines to facilitate our understanding of the invasion pathway and the features (organismal or environmental) that may assist in a species' invasion success (Kolar & Lodge 2001). Examples hereof include the intensity and frequency of propagule pressure (Lockwood et al. 2005; Colautti 2005; Simberloff 2009), dispersal mechanisms and pathways (Wilson et al. 2009) and genetic diversity (Lee 2002; Dlugosh & Parker 2008). In addition to these features, Blackburn et al. (2011) proposed a unified framework for biological invasions incorporating the stages of invasion, the potential barriers an invasive species needs to overcome during each stage to proceed to the next stage and the likely management strategies that may be employed to curb further invasion during each stage. Despite the progress made over the last 50 years, predictions and generalisations regarding a species' invasive success remain challenging and limited (García- Berthou 2007; Leprieur et al. 2009), likely due to the diversity of systems and taxa considered (Gaither et al. 2013).

Although general guidelines and unified frameworks are essential in furthering our understanding of biological invasions, exploring the genetic and phenotypic variation in a species and the influence of extrinsic factors during adaptation to novel environments could provide insight into a species' invasion success. For instance, a loss in genetic diversity is often associated with founder effects and population bottlenecks (Mayr 1963) experienced by the introduced alien species in the novel range upon colonisation (Eales et al. 2010). Lessons from conservation genetics suggest that this loss may limit a populations' ability to adapt in the novel

range, while the reduced population size simultaneously increases the risk of extinction (Allendorf & Lundquist 2003) and decreases the colonisation potential of the alien species (Reznick et al. 1997; Eales et al. 2010). However, founder effects may be countered via the presence of multiple genotypes within the introduced population, resulting from either multiple introductions (Kolbe et al. 2004) or the presence of an array of genotypes due to the admixed nature of the introduced population (Rius and Darling 2014). Beside genetic traits, rapid morphological adaptations (either phenotypic plasticity or local adaptation) to the novel environment have been suggested to play a key role in the successful establishment and subsequent spread of alien species in a novel environment (Agrawal 2001; Ghalambor et al. 2007; Lucek et al. 2014). By rapidly remodelling and subsequently adapting the phenotype to the new optimum, morphological adaptations are thought to shield the small introductory populations from strong selection (Lee 2002; West-Eberhard 2003; Cerwenka et al. 2014), allowing alien species survival and establishment in the novel environment.

South Africa, more specifically the Cape Fold Ecoregion (Abell et al. 2008; Ellender et al. 2017), is characterised by extensive freshwater fish diversity with high levels of endemism (Ellender et al. 2017). Ironically, the country is also recognised as a fish invasion hotspot (Leprieur et al. 2008), with 27 alien fish species being introduced into various freshwater ecosystems across the country (Ellender & Weyl 2014). Four of these species, namely *Micropterus salmoides* (largemouth bass), *Micropterus floridanus* (Florida largemouth bass), *Micropterus dolomieu* (smallmouth bass) and *Micropterus punctulatus* (spotted bass) belonging to the family Centrarchidae, have successfully established in South Africa (de Moor & Bruton 1988; Marr et al. 2017; www.invasives.org.za). While *M. salmoides* is recognised by the IUCN Global Invasive Species Database as one of the world's 100 worst invasive alien species (www.iucngisd.org), and has consequently received considerable attention with regard to its invasive potential, *M. dolomieu* (Lacepède 1802), in contrast, remains relatively data poor.

The freshwater genus *Micropterus* comprises eight recognised species and one subspecies (Figure 1.1; Lee et al. 1980; Bagley et al. 2011), all of which are endemic to North America (Lee et al. 1980). *Micropterus dolomieu*, in particular, is native to the southern parts of two Canadian provinces and 23 freshwater systems in the east-central parts of the United States of America, including the Great lakes (Loppnow et al. 2013). Currently recognised as an invasive species in at least 12 countries worldwide, *M. dolomieu* was initially introduced

into South Africa from the United States of America in 1937 for angling purposes (Loppnow et al. 2013). As a result of being a popular game fish around the world, most intentional and unintentional additional introductions have been caused by anglers (Skelton 2001; Loppnow et al. 2013).

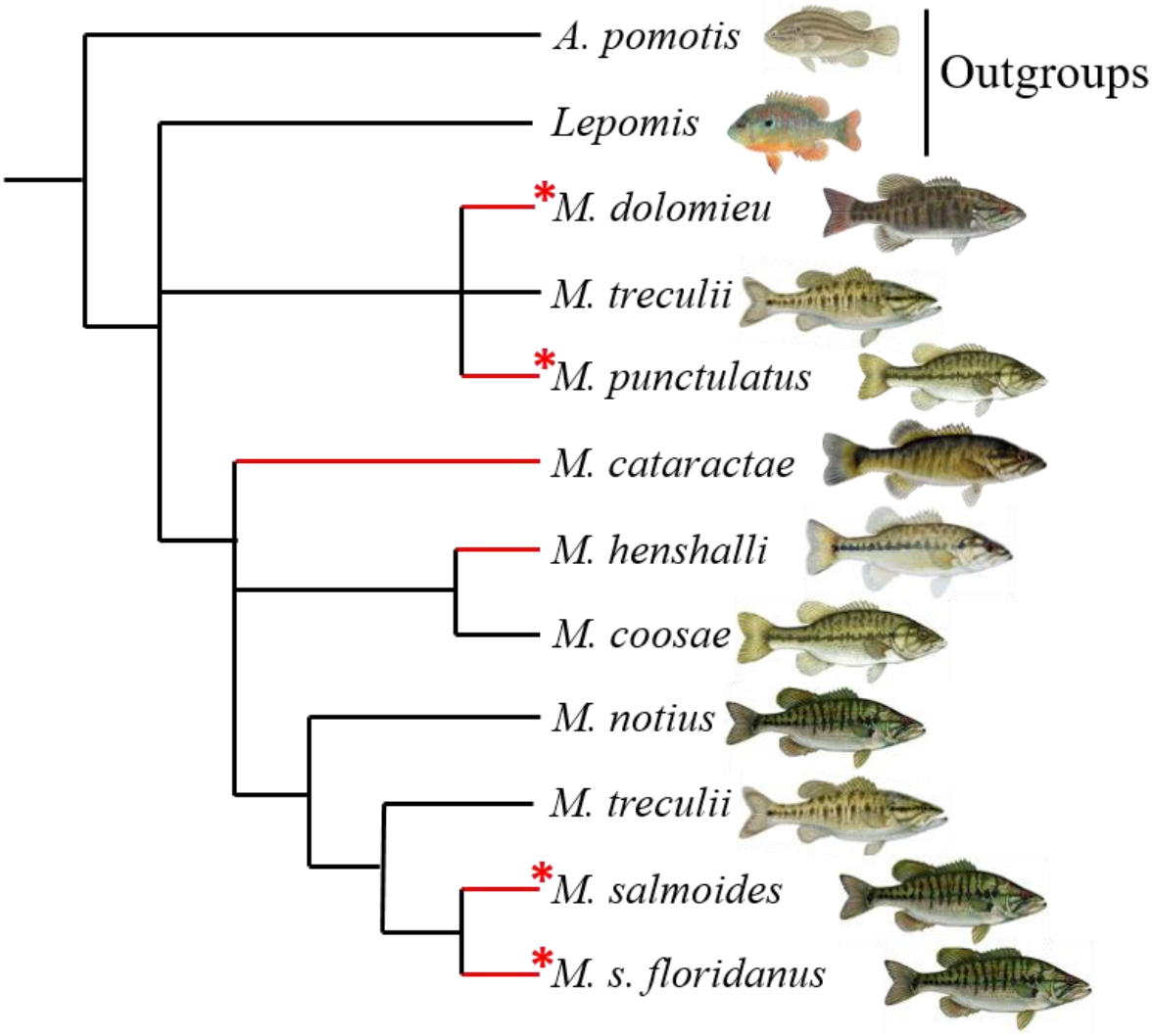


Figure 1.1. A reconstructed phylogeny of the recognised *Micropterus* taxa based on the results of Near et al. (2005) and Bagley et al. (2011). Red branches depict species that have invasive populations within North America while the red asterisks (*) represents the invasive *Micropterus* species found in South Africa.

Smallmouth bass are generalist littoral predators, feeding mostly on small fish and crustaceans (Skelton 2001; Loppnow et al. 2013). As ecosystem engineers, they can alter the novel invaded habitat both directly and indirectly through predation on native fish species, competing for food and nesting sites, and altering the food web structure (Ellender et al. 2011; Ricciardi & MacIsaac 2011). Despite being recognised as one of the most successful non-native invaders in the Western Cape of South Africa (de Moor & Bruton 1988; Van der Walt et al.

2016), relatively few studies (neither in its native or invasive range) have been conducted on the invasive potential and impact of *M. dolomieu*, with most drawing their conclusions from studies conducted on the sister species, *M. salmoides* (see Loppnow et al. 2013). It is, however, known that *M. dolomieu* can invade and establish itself in a range of both natural and artificial water bodies (impoundments, rivers, tributaries and streams) spanning an array of aquatic ecoregions (Figure 1.2; Skelton 2001) in a very short period of time (Vander Zanden et al. 2004; Stepien et al. 2007; Van der Walt et al. 2016). Due to the human facilitated movement of this species across the country to areas of recreational importance, *M. dolomieu* also spans a number of anthromes (Ellis & Ramankutty 2008). Furthermore, because *M. dolomieu* prefers flowing waters to stagnant pools, as is the case for *M. salmoides*, they are an ideal template taxon for elucidating the characteristics of invasiveness as their novel invaded range would likely be comprised of the full range of available habitats, ranging from natural headwater streams and mainstream rivers to artificial impoundments.

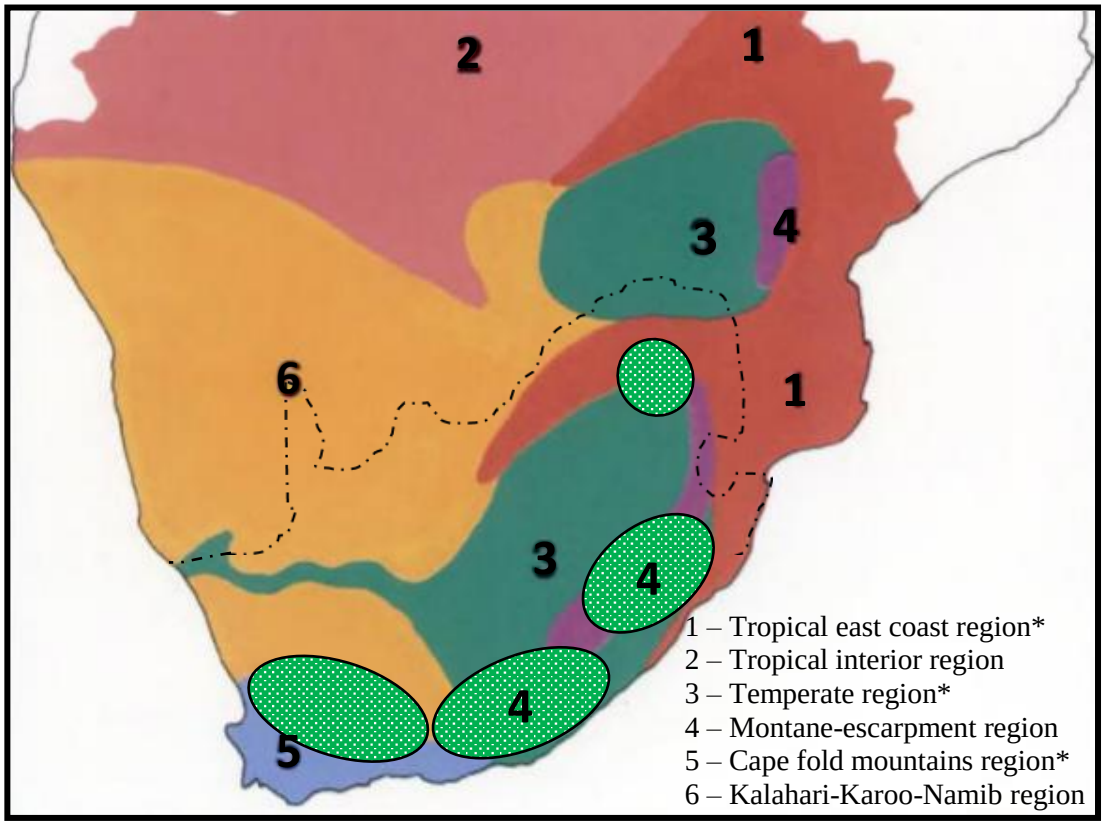


Figure 1.2. The six aquatic ecoregions of southern Africa as suggested by Skelton (2001). The ovals denote the distribution of *Micropterus dolomieu* within South Africa, as illustrated by the dashed outline. Known *M. dolomieu* populations can be found within ecoregion 1, 3 and 6, as indicated by the asterisk (*) in the legend. Figure adapted from Skelton (2001).

By utilising methods that span the disciplines of ecology, molecular biology, population genetics and functional morphology, I evaluated the potential mechanisms driving a species' invasive success using *M. dolomieu* as model organism. Firstly, I examined the temporal changes in genetic diversity by comparing historic and contemporary genetic data within the species' native and invaded South African range to discern patterns of genetic variation when an alien organism is released into a novel environment (Chapter 2). Next, I tested the hypothesis that hybridisation and introgression can occur between two invasive species in a novel invaded range (Chapter 3), hereby assisting the invasive species in question to adapt to less favourable environments during the early stages of invasion. In Chapter 4, I assessed the morphological variation, and tested the hypothesis that environmental variation drives morphological changes in phenotype. These morphological changes are predicted to be due to phenotypic plasticity, but as gene flow plays an important role in driving phenotypic variation among populations, the genetic basis to local adaptation is investigated in Chapter 5. The final chapter, Chapter 6, is dedicated to summarising the main findings of this thesis and attempts to extrapolate these findings to other freshwater invaders, suggesting key features and traits that should be investigated when looking at the invasive potential of other organisms.

CHAPTER 2

CHARACTERISING THE GENETIC DIVERSITY OF *MICROPTERUS DOLOMIEU* IN ITS NATIVE VERSUS AN INVASIVE RANGE: PERSPECTIVES FROM THE PAST AND PRESENT

ABSTRACT

Genetic diversity has long been thought essential to an introduced species' invasive success in the novel invaded range, with newly established populations often displaying low genetic diversity when compared to native populations. Reconstruction of the demographic history of an invasive species, by comparing genetic diversity levels across native and invasive ranges, is often used to delineate the most likely invasion scenario. Many studies, however, focus solely on contemporary samples, relying heavily on the premise that the historic population structure within the native range has been maintained over time. Instead, comparing historic and contemporary DNA may constitute a more powerful approach to detect recent demographic changes, as it allows for the monitoring of temporal changes in genetic diversity across generations. Using the invasive freshwater fish *Micropterus dolomieu*, introduced into South Africa in 1937, as model organism, I aim to (1) assess the genetic differentiation between- and genetic diversity within *M. dolomieu* populations in South Africa, (2) unravel how the genetic diversity changed over time in both native and invasive ranges, and (3) confirm the single historic introductory event for *M. dolomieu* into South Africa. A three-way comparison conducted on two mitochondrial- and nine microsatellite loci for a total of 572 *M. dolomieu* specimens, representing the contemporary invasive South African range, contemporary native USA range and the historic native USA range (dating back to the period of introduction into South Africa), revealed elevated levels of genetic diversity for the invasive South African range. The levels of genetic diversity for the contemporary native and invasive ranges were, however, substantially lower than those of the historic native range, suggesting that both contemporary populations experienced a recent genetic bottleneck. Furthermore, the invasive South African range displayed significant levels of population structure, while both historic and contemporary native USA populations revealed higher levels of admixture. Comparison of contemporary and historical samples revealed both a historic introduction of *M. dolomieu*, as well as a more recent introduction, thereby disproving the notion of a single introduction into the invaded South African range. Although multiple introductions might have

contributed to the high levels of genetic diversity in the invaded range, I discuss alternative factors that may have been responsible for the elevated levels of genetic diversity.

INTRODUCTION

Genetic diversity is essential to the invasive success of an introduced species in the novel invaded range (Kolbe et al. 2004; Gillis et al. 2009; Funk et al. 2011; Beneteau et al. 2012; Rius & Darling 2014).It is generally assumed that a newly established population will have lower genetic diversity when compared to populations in the native range because of the small founding population size and genetic bottleneck experienced on arrival in the novel range (Mayr 1963; Nei et al. 1975; Roman & Darling 2007; Eales et al 2010, but see Lee et al. 2004; Novak & Mack 2005; Wares et al. 2005). The low genetic diversity is expected to have detrimental effects on the colonisation potential of the alien species (Reznick et al. 1997; Kolbe et al. 2004; Ficetola et al. 2008; Eales et al. 2010; Dormontt et al. 2014), ultimately leading to a decrease in population fitness and adaptability to the invaded range (Nei et al. 1975; Keller & Waller 2002; Guinard et al. 2003; Rius & Darling 2014). Indeed, a meta-analysis by Dlugosch & Parker (2008) on 80 plant, animal, and fungi species, revealed that genetic diversity does decrease during an invasion. However, the decrease in genetic diversity can be attributed to the loss in allelic diversity rather than the loss of heterozygosity (Dlugosch & Parker, 2008). Heterozygosity, defined as the number of heterozygous individuals at a locus, and allelic diversity (i.e. allelic richness), defined as the number of alleles representing a locus, are often used to measure genetic variation (Allendorf 1986). Though the former measure has been used more extensively than the latter, heterozygosity measures are also relatively insensitive to the real number of genotypes present at a particular locus (Allendorf 1986; Petit et al. 1998). This is because rare alleles are often not accounted for when calculating heterozygosity measures, as explained by Allendorf (1986). Hence, allelic richness is thought to be a better measure for portraying a populations' long-term evolutionary potential (Allendorf 1986; Petit et al. 1998; Leberg 2002).

The loss of genetic diversity in the novel invaded range can, however, be circumvented through two non-mutually exclusive mechanisms, namely propagule pressure and admixture (Dlugosch & Parker 2008; Wilson et al. 2009; Dlugosch et al. 2015; Naccarato et al. 2015). On the one hand, propagule pressure (Simberloff 2009; Wilson et al. 2009) can increase genetic diversity through either multiple introductions or a single introduction comprising many individuals (Kelly et al. 2006; Dlugosh & Parker 2008; Simberloff 2009; Bouchard et al. 2011). On the other hand, admixture via hybridisation (both intra- and interspecific) allows novel allelic combinations to be generated (Kelly et al. 2006; Rius & Darling 2014). This, in turn, leads to an increase in short-term population fitness via heterosis, otherwise known as hybrid vigour (i.e. hybrid offspring displaying phenotypic superiority over parents) as well as increased adaptive potential (Rius & Darling 2014; Dlugosh et al. 2015), ultimately facilitating the species' invasion success (Kelly et al. 2006; Dlugosh et al. 2015). While some species maintain their genetic diversity levels through these mechanisms, others display elevated levels of genetic diversity when compared to the native range (Roman & Darling 2007; Dlugosch & Parker 2008), as observed in the brown anole, *Anolis sagrei* (Kolbe et al. 2004), the nassariid gastropod, *Cyclope neritea* (Simon-Bouhet et al. 2006), canarygrass, *Phalaris arundinacea* (Lavergne & Molofsky 2007) and the oriental shrimp, *Palaemon macrodactylus* (Lejeusne et al. 2014). Nevertheless, admixture may also be disadvantageous, erasing locally adapted phenotypes in established non-native populations (Lynch 1991; Lenormand 2002; Dlugosh et al. 2015) and reducing offspring viability due to genetic incompatibilities (Johansen-Morris & Latta 2006). Admixture is most often attained through multiple introductions, as seen in an array of studies (Kolbe et al. 2004; Yonekura et al. 2007; Gillis et al. 2009; Pairon et al. 2010; Funk et al. 2011; Beneteau et al. 2012; Rius & Darling 2014), but depending on the origin of these introductory populations, a variety of outcomes are plausible. For instance, Dlugosh et al. (2015) concluded that multiple introductions had little effect on genetic variation in the novel environment. However, if the introduced populations originated from different areas within the native range and admixture were to occur between these divergent source populations, multiple introductions could significantly impact the genetic variation within the invasive populations (Dlugosh et al. 2015).

Numerous studies have compared the molecular signature of invasive species in their native and invasive ranges (Guinard et al. 2003; Kolbe et al. 2004; Kelly et al. 2006; Rollins et al. 2009; Kawamura et al. 2010; Naccarato et al. 2015), in the hopes of unravelling the demographic history of the invasive populations (Ficetola et al. 2008; Gillis et al. 2009; Neilson

& Stepien 2011; Gray et al. 2014). However, most studies to date are restricted to contemporary specimens, thereby relying heavily on the premise that the historic population structure within the native range has been maintained over time and no allele frequency shifts have occurred. Historic DNA serves as a valuable reference when examining contemporary genetic diversity (Bouzat 2000; Guinard et al. 2003; Lozier & Cameron 2009; Dormontt et al. 2014), as it allows for the monitoring of temporal changes in genetic diversity across generations (Guinard et al. 2003; Sefc et al. 2007), while increasing the chance of detecting subtle changes frequently overlooked by studies focussing only on contemporary data (Lozier & Cameron 2009). Hence, despite the difficulties relating to the amplification and genotyping of historical samples (Sefc et al. 2003; 2007), comparing historic and contemporary DNA constitutes a powerful approach to detect recent demographic changes in invasive species, as it can delineate the most likely invasion scenario (Gillis et al. 2009; van Kleunen et al. 2010; Thompson et al. 2011) and reveal the connectivity between invasive populations (Funk et al. 2011; Beneteau et al. 2012).

The invasive smallmouth bass, *Micropterus dolomieu* (Lacepède, 1802), presents an ideal model system to investigate variation in genetic diversity through space and time, as its introduction history and subsequent spread into and throughout South Africa is well recorded (de Moor & Bruton 1988). Twenty-nine *M. dolomieu* specimens originating from broodstock collected in the Wheeling River, West Virginia, USA were shipped from the Lewistown hatchery in Maryland, USA, to the Jonkershoek hatchery in South Africa in 1937 (Powell 1967; de Moor & Burton 1988; Loppnow et al. 2013). Here, they were reared and bred before being released into multiple water bodies across the country for angling purposes (de Moor & Bruton 1988). Most of the intended stockings occurred between 1938 and 1981, starting with the Berg (1938-1939), Breede (1939-1940), Olifants (1943, 1945) and Buffalo (1949) River systems (Table 2.1; Figure 2.2), after which the remaining rivers and water bodies throughout South Africa were stocked (de Moor & Bruton 1988). The number of fingerlings initially released into the rivers was considerable, with more than 300 fingerlings reportedly released into the Berg River alone during the first introductory event in 1938 (de Moor & Bruton 1988).

In this chapter I will examine the genetic structuring within and between contemporary native (USA) and contemporary invasive (South African) *M. dolomieu* populations. Furthermore, using genomic DNA obtained from formalin-fixed muscle tissue, I will investigate how the genetic diversity within the native range changed over time (i.e. historic versus contemporary specimens) and how it compares to contemporary invasive South African

M. dolomieu populations. My aims are to (1) assess the genetic differentiation and diversity within *M. dolomieu* populations in South Africa, (2) unravel how the genetic diversity changed over time in both native and invasive ranges, (3) confirm the single historic introductory event for *M. dolomieu*. Given the small *M. dolomieu* founding population, I predict that the invasive South African range will have a lower genetic diversity when compared to the native (historic and contemporary) North American range (Figure 2.1) due to a loss of alleles, as suggested by Dlugosch & Parker (2008). Furthermore, as microsatellite loci have high mutation rates (Webster & Reichart 2005), and human-mediated dispersal has increased the native distribution range of *M. dolomieu* (Loppnow et al. 2013), new alleles may be created faster than lost due to drift. Thus, I predict that the genetic diversity would be higher in contemporary time when compared to historic samples in the native range (Figure 2.1).

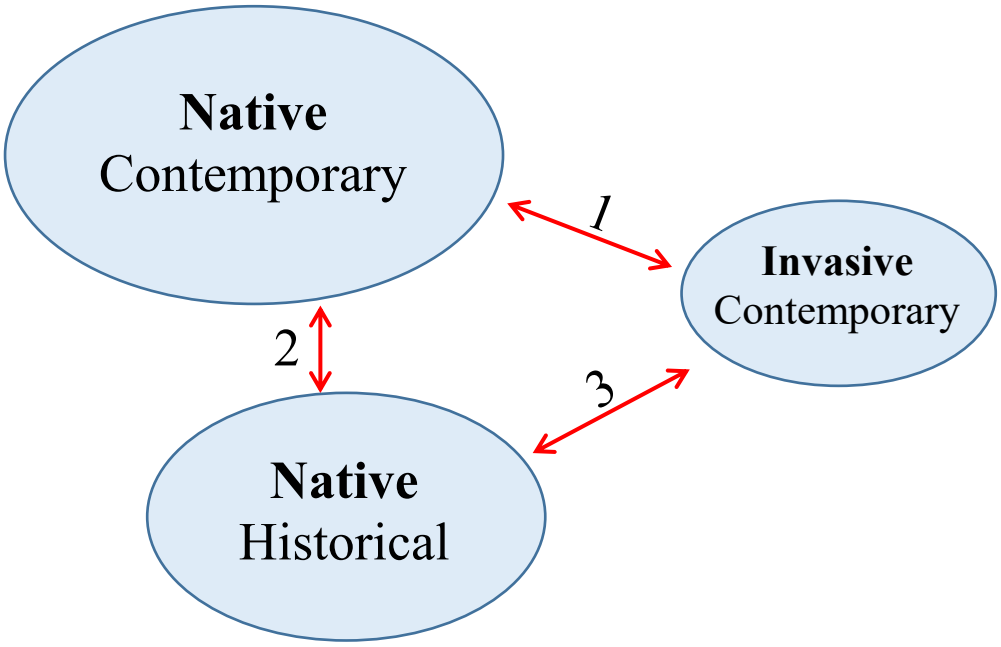


Figure 2.1. A schematic representation of the predictions made in this chapter. The size of each oval corresponds to the expected genetic diversity levels in each population, while the arrows represent the comparisons made in this study.

MATERIALS AND METHODS

DNA collection and extraction from historic specimens

Specimens representing the historic native range (Figure 2.2), corresponding to the approximate time of introduction into South Africa (1930 – 1941) were obtained from a host of collections housed at the Smithsonian National Museum of Natural History (NMNH), The Academy of Natural Sciences of Drexel University (ANSP), University of Michigan Museum of Zoology (UMMZ), and the Ohio State University Museum (OSUM) (Table 2.1; Appendix 2.1). In total, 53 formalin-fixed specimens representing 11 drainage systems, were obtained for genetic analysis (Table 2.1). These specimens represent a subset of the *M. dolomieu* genetic diversity that was present in the native range 20 – 25 generations ago (*M. dolomieu* usually require 4-5 years to reach sexual maturity, but elevated levels (i.e. 3 years) have been observed in invaded riverine systems; Bartel et al. 2008).

Genomic DNA was extracted from 20-50 mg preserved muscle tissue in a room previously unexposed to fish DNA. Prior to each extraction, all equipment and surfaces were treated with 10 % bleach to remove any potential contaminants. Where possible, each sample was carefully sub-sampled and placed in a 1.5 ml Eppendorf tube. Pikor et al. (2011) showed that high quality DNA can be extracted from formalin-fixed tissue if the samples are rehydrated with a series of ethanol washes prior to extraction. Thus, 500 µL of 100 % ethanol was added to each tissue sample in the Eppendorf tube and vortexed vigorously for 30 seconds. The liquid was removed and the process was repeated with 500 µL 70 % ethanol, followed by 1000 µL distilled water. Lastly, 1000 µL distilled water was added to each Eppendorf tube and left to soak at 55 °C for five days, vortexing the sample every 24 hours. Once rehydrated, the sample was moved to a dry Eppendorf tube before starting with the DNA extraction, using the commercially available QIAamp DNA FFPE tissue extraction kit (QIAGEN). In a recent review by Paireder et al. (2013) it was demonstrated that this kit consistently outcompeted other extraction methods when working with old (1820 - 1950), formalin fixed tissue. Apart from doubling the amount of Proteinase K added to each sample (60 µL), extraction followed the manufacturers' protocol. To break the formalin bonds, the samples were heated to 90 °C for one hour before commencing with the wash steps. As DNA denatures at 94 °C, the temperature was monitored and adjusted if necessary to avoid accidental overheating of samples. Lastly, to ensure the maximum elution of bound DNA, 10 µL elution buffer (warmed to 25.5 °C) was

dispensed onto the centre of the spin column membrane and left to ‘incubate’ at room temperature for 5 minutes before centrifuging at 14, 000 rpm for 1.5 minutes. This was repeated three times to yield a total DNA extraction volume of 30 µL. All DNA extractions were stored at -20 °C.

Table 2.1 An overview of the sampled populations from the native historic, native contemporary and invasive contemporary range. Abbreviations correspond to those used in subsequent tables, text and Appendix 2.1.

	Native/ Invasive	State/ Province	Collection Date	Sampled Locality	Drainage System	Abbr. in Tables	n	Formaldehyde Exposure	Material Supplied By	Symbol on Sampling Map (Fig. 2.2)
Historic Specimens	Native	Ohio	1930	White Oak Creek	Ohio River	OH	3	yes	OSUM	A
	Native	Ohio	1940; 1941	Auglaize River	Auglaize River	AU	5	yes	OSUM	B
	Native	Michigan; Ontario	1934; 1935; 1940	Detroit River	Detroit River	DET	18	yes	UMMZ	C
	Native	Ohio	1941	Lake Erie	Lake Erie	LE	3	yes	OSUM	D
	Native	Ohio	1938	Mosquito Creek Lake	Mosquito Creek	MO	2	yes	OSUM	E
	Native	New York	1937	Allegheny River	Allegheny River	AL	3	yes	UMMZ	F
	Native	New York	1931	Fall Creek	Cayuga Lake, Etna	FC	2	yes	UMMZ	G
	Native	New York	1935	Otselic River; Susquehanna River	Susquehanna River	SU	5	yes	UMMZ	H
	Native	New York	1936	Rondout River	Hudson River	HUD	4	yes	UMMZ	I
	Native	Maryland	1941	Monocacy River	Potomac River	PO	4	no	ANSP	J
	Native	Virginia; West Virginia	1933 - 1936	Shenandoah River	Shenandoah River	SH	4	yes	NMNH	K
							53			
Contemporary Specimens	Native	Ontario	2013; 2014	Detroit River	Detroit River	DET	7	yes	ROM	1
	Native	New York	2014	Niagra River	Niagra River	NIA	49	no	USA collectors	2
	Native	New York	2014	St Lawrence River	St Lawrence River	STL	55	no	USA collectors	3
	Native	New York	2015	Oneida Lake	Oneida River	ONEI	27	no	USA collectors	4
	Native	New York	2015	Saratoga Lake	Hudson River	SAR	10	no	USA collectors	5
	Native	New York	2015	Vestal; Susquehanna River	Susquehanna River	VES	14	no	USA collectors	6
	Native	New York	2015	Oneonta; Susquehanna River	Susquehanna River	ONEO	10	no	USA collectors	7
	Native	New York	2015	Lolliersville	Susquehanna River	LOL	20	no	USA collectors	8
	Native	New York	2014	Hudson River	Hudson River	HUD	21	no	USA collectors	9
							213			
	Invasive	Western Cape	2014	Doring River	Doring River	DO	38	no	Self-Collected	1
	Invasive	Western Cape	2014	Olifants River; Jan Dissels River	Olifants River	OL	44	no	Self-Collected	2
	Invasive	Western Cape	2014	Berg River	Berg River	BE	22	no	Self-Collected	3
	Invasive	Western Cape	2014	Breede River	Breede River	BR	43	no	Self-Collected	4
	Invasive	Eastern Cape	2014	Kouga River	Kouga River	KO	46	no	Self-Collected	5
	Invasive	Eastern Cape	2012	Krom River	Krom River	KR	15	no	SAIAB	6
	Invasive	Eastern Cape	2014	Rooikranz Dam	Buffalo River	BU	48	no	SAIAB	7
	Invasive	Mpumalanga	2014	Blyde Dam	Blyde River	MP	50	no	MPB	8
							306			

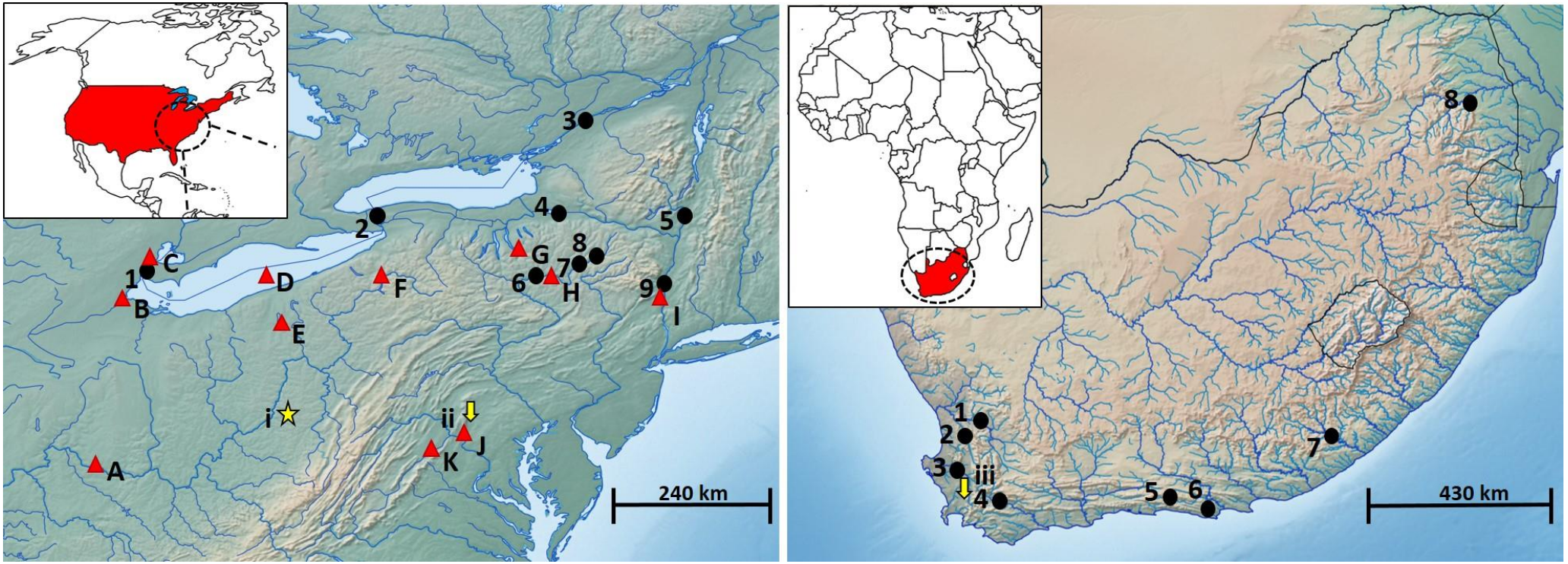


Figure 2.2. Map of native USA (left) and invasive SA (right) sampling localities. Letters A-K denote historic sampling localities, while numbers denote contemporary sampling localities. All letters and numbers correspond to those used in Table 2.1. The location indicated by the star (i) represent the Wheeling River, while the downward-facing arrows denote the (ii) Lewistown hatchery and (iii) Jonkershoek hatchery, respectively.

DNA collection and extraction from contemporary specimens

Fresh tissue samples (muscle, liver, fin clippings) were collected via angling methods in both the native United States of America (USA) and the invasive South African (SA) ranges during the respective summer months of 2014 and 2015 (Figure 2.2). Collections in the USA were led by a host of individuals and organisations based in the USA (see acknowledgements). Nine localities, rendering a total of 213 specimens were sampled from the same 'broad' area represented by the historic samples to allow for direct genetic diversity comparisons (Table 2.1). As only fin clippings were collected in the USA, all tissue samples were shipped either dry in paper envelopes or in 70 % ethanol. Addition specimens collected in 2014 (n = 7; formalin fixed), representing the Detroit River, were obtained from the Royal Ontario Museum (ROM), Canada.

All SA specimens were euthanised with clove oil (CapeNature permit number 0056-AAA043-00004; Eastern Cape permit numbers CRO 165/14CR and CRO 166/14CR; Mpumalanga permit number MPB. 5498/2; Ethical clearance reference number SU-ACUM14-00011, University of Stellenbosch) upon capture, before sampling a 20-50 mg piece of muscle/ liver tissue. Tissue samples were stored in 70 % ethanol for further DNA analysis. Additional specimens were obtained from the South African Institute for Aquatic Studies (SAIAB), Grahamstown, South Africa, rendering a total sample size of 306 specimens representing eight river systems (Table 2.1; Appendix 2.1). DNA was extracted from each contemporary specimen (USA and SA) using the NucleoSpin Tissue extraction (gDNA) kit (MACHEREY-NAGEL, Separations, Cape Town, South Africa) following the manufacturers protocol. All DNA extractions were stored at -20 °C.

Historic and contemporary DNA amplification

Two partial mitochondrial gene regions, namely cytochrome b (cytb) and control region (CR) were amplified for all the contemporary samples (n = 519) using a polymerase chain reaction (PCR). However, due to the degraded nature of the historic DNA, the limited availability of tissue and the low amount of DNA following extraction, mtDNA amplification was not possible. A standard 25 µL mastermix was prepared for both mtDNA PCR reactions, and contained the following: 01 µL Supertherm Taq polymerase, 2.5 µL buffer, 3 µL $MgCl_2$ (all three supplied by JMR Holdings), 0.5 µL dNTP mix (10 mM), 0.5 µL each of forward and reverse primers (10 mM), 14.9 µL dH2O and 2-5 µL genomic DNA. The internal cytb primers, basscytbf1 (5´-CAC CCC TAC TTC TCC TAC AAA GA- 3´) and basscytbr1 (5´-AAG GCR

AAG CGG GTG AGG G- 3′; Near et al. 2003) were used to amplify the cytb fragment under the following PCR thermocycling profile: denaturation at 95 °C for 1 minute, followed by 40 cycles of 40 seconds at 95 °C, 40 seconds at 56 °C (annealing temperature) and 1 minute at 72 °C, with the final extension performed for 10 minutes at 72 °C. Primer set CB3R-L (5 ′ - CATATTAAACCCGAATGATATTT- 3′; Palumbi 1996) and HN20-R (5′ - GTGCTTATGCTTTAGTTAAGC- 3′; Bernatchez & Danzmann 1993) were used to amplify the CR. PCR conditions for CR were identical to those used to amplify cytb, with the exception of the annealing temperature, which was set to 50 °C. Despite numerous attempts to amplify the seven contemporary Detroit River (DET) samples for CR, no amplification was achieved, most likely due to the formalin fixation used to preserve the tissue. For all PCR reactions, positive and negative controls were run and all amplified PCR products were visualised through gel electrophoresis before being sequenced (ABI 3730 XL DNA Analyzer, Applied Biosystems, CAF, Stellenbosch, South Africa). Chromatographs were visually inspected and aligned in Geneious® 10.0.2 (Biomatters, Auckland, New Zealand).

Fifteen microsatellite loci, designed for both species- and genus-level amplification, were selected from published literature (Table 2.2). Of these, only 11 microsatellite loci (eight species-specific: Mdo3, Mdo4, Mdo5, Mdo7, Mdo8, Mdo9, Mdo10, Mdo11 - Malloy et al. 2000; three genus-specific: Lma21 - Colbourne et al. 1996; Lma102, Lma117 - Neff et al. 1999) were successfully amplified. Lma102 and Lma117 were not polymorphic for a subset of specimens and were therefore excluded from the study. Hence, nine polymorphic loci were used in the present study (Table 2.2). Three multiplex reactions were used to amplify the nine microsatellites, using the KAPA2G™ Fast Multiplex PCR Kit (KapaBiosystems, Cape Town, South Africa). Each 10 µL reaction comprised the following: 5 µL Kapa2G™ Fast Multiplex Mix, 0.2 µL of each primer (10 nmol) and 1 µL template DNA. The PCR cycling parameters followed those specified under the Low Plex cycling step, as specified by Kapa2G™ Fast Multiplex PCR Kit: denaturation for 3 min at 95 °C, followed by 30 cycles of 15 s at 95 °C, 30 s at primer dependant annealing temperature (see Table 2.2) and 30 s at 72 °C, with the final extension being performed for 10 min at 72 °C.

To minimise the potential for PCR contamination of the historic samples, all molecular work was conducted with a set of pipettes not previously used to amplify fish DNA. The same nine microsatellite loci were amplified for the historic samples and followed the same amplification procedure used for the contemporary DNA. Similarly, three multiplex reactions

were conducted using 3 µL pure DNA in each, but due to the degraded nature of the DNA this did not yield any results. Thus, the resulting PCR products for each multiplex were then diluted with distilled water to obtain a $^{1}/_{10}$ PCR product dilution which, in turn, served as the DNA in the subsequent PCR. To ensure accurate amplification and to try and avoid the overestimation of genetic diversity often associated with the amplification of ancient- and/ or formalin-fixed DNA (Buchan et al. 2005; Sefc et al. 2007), historic samples were amplified twice for each microsatellite locus. All microsatellite genotyping (contemporary and historic samples) was performed on the ABI 3730 XL DNA Analyzer (Applied Biosystems, CAF, Stellenbosch, South Africa), using LIZ as an internal size marker. To ensure accurate scoring, reference individuals previously scored were used as positive controls. Geneious® 10.0.2 (Biomatters, Auckland, New Zealand) was used to score the microsatellites. Historic specimens were scored blindly (i.e. all specimen names were 'hidden' as to prevent any bias), and repeated three times to ensure accuracy. Where scoring inconsistencies were observed (historic specimens) and/ or more than three loci could not be scored (for both historic and contemporary specimens), the entire specimen was removed from the dataset and excluded from the study. Similarly, since one microsatellite locus, namely Mdo8, did not amplify for the majority of historic samples, it was removed from the historic dataset entirely. Thus, nine microsatellite loci were analysed for the contemporary dataset, but only eight microsatellite loci were analysed for the historic dataset.

Table 2.2 Microsatellite loci amplified in the present study with the corresponding primer sequence, reference, size, repetition pattern, optimised annealing temperature, multiplex reaction each locus was part of and dye labels used. Shaded areas represent microsatellite loci that were excluded from the study due to amplification errors.

Locus	Primer Pair and Sequence	Reference	Size (bp)	Core Repeat	Amplified	Multiplex #	Ta (°C)	Dye Label
Mdo 1	F: 5' GCTCTTCCCAGTGGTGAGTC 3' R: 5' ATCTCAGCCCATACCGTCAC 3'	Malloy et al. 2000	210	$(GT)_{14}$	X			
Mdo 2	F: 5' GCCCTTTCATATTGGGACAA 3' R: 5' CTGCTCTGGCGTACATTTCA 3'	Malloy et al. 2000	197	$(GT)_{14}$	X			
Mdo 3	F: 5' AGGTGCTTTGCGCTACAAGT 3' R: 5' CTGCATGGCTGTTATGTTGG 3'	Malloy et al. 2000	135	$(CA)_{20}$	√	1	53.9	6-FAM
Mdo 4	F: 5' TCTGAACAACTGCATTTAGACTG 3' R: 5' CTAATCCCAGGGCAAGACTG 3'	Malloy et al. 2000	142	$(CA)_{11}$	√	1	53.9	NED
Mdo 5	F: 5' CAGGTTCCCTCTCACCTTCA 3' R: 5' ATGGTCTCACCAGGGACAAA 3'	Malloy et al. 2000	200	$(CT)_9CC(CA)_{10}GA(CA)_3TA(CA)_2$	√	2	61.0	PET
Mdo 6	F: 5' TGAAATGTACGCCAGAGCAG 3' R: 5' TGTGTGGGTGTTTATGTGGG 3'	Malloy et al. 2000	150	$(CA)_7(TA)_4$	X			
Mdo 7	F: 5' TCAAACGCACCTTCACTGAC 3' R: 5' GTCACTCCCATCATGCTCCT 3'	Malloy et al. 2000	172	$(CA)_{12}$	√	1	53.9	VIC
Mdo 8	F: 5' GTGAGGACCAGCCAAAATGT 3' R: 5' GGAAGATTGAGGTCCCAACA 3'	Malloy et al. 2000	220	$(CA)_{19}$	√	3	58.3	NED
Mdo 9	F: 5' TTTGATGGGCGTTTTGTGTA 3' R: 5' GACCGGTCCTGCATATGATT 3'	Malloy et al. 2000	126	$(GT)_{10}$	√	3	58.3	PET
Mdo 10	F: 5' GTGTCTCCGTGTGTTGATGG 3' R: 5' AACACCAGAGGCAAACAAGC 3'	Malloy et al. 2000	101	$(GT)_{10}$	√	3	58.3	VIC
Mdo 11	F: 5' TTGTGGAGAGGGGCATAAAC 3' R: 5' GCATCCTCCCACGTTACCTA 3'	Malloy et al. 2000	174	$(GT)_{11}GA(GT)_3$	√	3	58.3	6-FAM
Lma21	F: 5' CAGCTCAATAGTTCTGTCAGG 3' R: 5' ACTACTGCTGAAGATATTGTAG 3'	Colbourne et al. 1996	158-182	$(TC)_{19}(AC)_{11}$	√	2	61.0	6-FAM
Lma87	F: 5' ATGACACAGACTCACCATGC 3' R: 5' CTCCTGCCCATAAATCAGAC 3'	Colbourne et al. 1996	118-152	$(AC)_{15}A_5$	X			
Lma102*	F: 5' CTGTGAAAATGGTGTGAGCG 3' R: 5' AAACACAAAAGTCCACGCAC 3'	Neff et al. 1999	88-102	$(GT)_{19}ATGTAT(GT)_4$	√	2	61.0	NED
Lma117*	F: 5' CCACCAACAGCATGCAGAC 3' R: 5' CATGCCACTCATTGCACTG 3'	Neff et al. 1999	194-218	$(GT)_{22}$	√	2	61.0	VIC

* Amplified microsatellite loci, but not polymorphic. Hence, both Lma102 and Lma117 were excluded from the dataset.

Contemporary mtDNA analyses

To assess the diversity levels in both the native (USA) and invasive (SA) ranges, the number of haplotypes, haplotype diversity and nucleotide diversity (*H*, *h* and π, respectively), were calculated in ARLEQUIN 3.5.2.2 (Excoffier & Lischer 2010). The population history for *M. dolomieu* in both ranges was examined using Fu's *Fs* (Fu, 1997) and Tajima's *D* (Tajima 1989), assessing each with 10 000 permutations as implemented in ARLEQUIN 3.5.2.2 (Excoffier & Lischer 2010).

To assess the genetic population structure, both native and invasive contemporary datasets were combined for each gene fragment. A parsimonious haplotype network was constructed for both cytb and CR, using the 95 % connection limit as implemented in TCS 1.21 (Clement et al. 2000). Gaps and/ or missing data were treated as a fifth state. Estimation of Wright's pairwise F_{ST}, and a hierarchical Analysis of Molecular Variance (AMOVA) was conducted to assess the genetic population structure among sampled localities. Analyses were conducted in ARLEQUIN 3.5.2.2 (Excoffier & Lischer 2010), with statistical significance assessed after 10 000 permutations. Wright's pairwise F_{ST} measures the amount of population subdivision, with values near to 0 indicating no population differentiation, and values near to 1 indicate complete population differentiation (Excoffier et al. 2005), while a hierarchical AMOVA assesses the degree of molecular variation within versus among sampled localities.

Contemporary and historic microsatellite analyses

All genotype frequencies were assessed for linkage disequilibrium (i.e. the non-random combination of alleles between loci) and deviations from Hardy-Weinberg equilibrium (HWE), using Genepop 4.2.1 (Rousset, 2008), with statistical significance being assessed after 10 000 iterations. The Bonferroni method was used to correct for multiple comparisons (Rice 1989). Amplification errors associated with large allele drop-out and stuttering was assessed with MICROCHECKER (Van Oosterhout et al. 2006). As most of the populations were found to not comply with HWE assumptions, FreeNA (Chapuis & Estoup 2007) was used to check for the presence of null alleles using the EM algorithm (Dempster et al. 1977). Following Chapuis & Estoup (2007), null allele frequency values above 25 % signified the presence of null alleles, while null allele values below 25 % were thought to not significantly impact the results. Intraspecific and within-population genetic diversity levels were assessed as number of alleles (Na), allelic richness (AR) observed heterozygosity (H_O), expected heterozygosity (H_E), and Wright's inbreeding coefficient (F_{IS}), as implemented in FSTAT (Goudet 1995), Genepop (Rousset 2008), HP-Rare (Kalinowski 2005) and ARLEQUIN (Excoffier & Lischer 2010).

Statistical significance of F_{IS} was assessed after 1000 permutations in FSTAT (Goudet 1995). F_{IS} ranges from -1 (excess number of heterozygotes) to 1 (excess number of homozygotes), with populations displaying significantly positive F_{IS} values experiencing high levels of inbreeding (Beebee & Rowe 2005). Allelic richness (AR) was calculated using HP-Rare (Kalinowski 2005), as this program corrects for sample size disparity through rarefaction analysis. All analyses were conducted per population for the two contemporary datasets, but due to the small sample size for most of the historic localities (Table 2.1) all localities were grouped to obtain the genetic diversity indices.

Multiple approaches were employed to investigate the population structuring and genetic connectivity among (contemporary and historic) populations. As only eight loci were amplified for the historic specimens, all comparative analyses incorporating the historic samples only compared the eight loci, while contemporary SA – USA comparisons encompassed all nine loci. Firstly, to determine whether there is a difference in observed heterozygosity (H_O) between the three groups (contemporary invasive – C.I, contemporary native – C.N, historic native – H.N), an Analysis of Variance (ANOVA) was conducted in SPSS STATISTICS v. 20.0.0 (SPSS Inc., Chicago, IL, USA), with loci selected as random factor. Subsequently, a Bonferroni post-hoc test was used to further assess the differences between groups. In addition, a stacked bar graph was constructed to visualise the variation among localities and loci. Secondly, Weir's (1986) F_{ST} was employed to assess the genetic differentiation among sampled localities using FreeNA (Chapuis & Estoup 2007). FreeNA was chosen, employing the *ENA* correction method (Chapuis & Estoup 2007), as this method has been shown to correctly estimate F_{ST} values in the presence of null alleles (detected in the previous analysis) (Chapuis & Estoup 2007). A jackknife approach (Chapuis & Estoup 2007) implementing 1000 bootstrap replicates, was employed to assess the statistical significance. F_{ST} values can range between 0 to 1, with low values signifying no differentiation and higher values suggesting complete differentiation. Genetic differentiation is thought to be high when values exceed 0.2, while values near to 0 suggest undifferentiated populations (Excoffier et al. 2005).

To investigate the genetic associations within each of the three groups (C.I, C.N, H.N) as well as among them, without being influenced by the lack of HWE or the presence of null alleles, a Principal Component Analysis (PCA) using microsatellite allelic frequencies, was conducted in the R package Adegenet 1.3.1 (Jombart & Ahmed 2011). Next, a Bayesian

clustering approach was used to (a) identify and visualise the population structure within each of the three groups, (b) search for a potential source population from where the invasive South African stocks originated, (c) compare overlapping populations from the historic- and contemporary native range to determine the degree of genetic similarity, using STRUCTURE version 2.3.4 (Pritchard et al. 2000). STRUCTURE, assuming HWE, assigns each sampled individual to a cluster without prior knowledge of the samples' geographical origin, ultimately aiming to identify the most probable number of genetic clusters (*K*). All four STRUCTURE analyses (each group independently followed by an analysis combining C.I, C.N and H.N) were conducted using the admixture model with correlated allele frequencies, allowing each individual to be allocated to multiple clusters, as determined by its genotype frequency. Five replicate runs were conducted for each *K*, with *K* = 1-15 for each independent analysis. An initial burn-in of 75 000 Markov Chain Monte Carlo (MCMC) generations were run, followed by 350 000 MCMC iterations. Following the Evanno method (Evanno et al. 2005), STRUCTURE HARVESTER (Earl & vonHoldt 2012) was used to determine the most probable *K*, before using CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) to compile the five replicate runs for the most likely *K*. DISTRUCT 1.1 (Rosenberg 2004) was used to visualise the composite assignments.

Lastly, to determine whether the invasive South African *M. dolomieu* populations originate from a single introductory event from the USA, an Approximate Bayesian Computation (ABC) was performed on the microsatellite dataset with DIYABC v2.1.0 (Cornuet et al. 2014). Once again, the sampled localities (Table 2.1) were pooled into three groups, i.e. C.I, C.N and H.N. Six simple yet competing introduction scenarios were generated under a coalescent framework (Figure 2.7: 1 - 6) in order to focus the computational efforts on probable introduction scenarios rather than an exhaustive list of possibilities. Scenario 1: C.I originated from the H.N stock which represents a subsample of the C.N populations; Scenario 2: C.I originated from C.N populations, with both populations being derived from H.N (i.e. a more recent introduction event than the one on record); Scenario 3: C.I did not originate from either C.N or H.N population, but rather from an unsampled population; Scenario 4: C.I populations represent admixed populations from both C.N and H.N; Scenario 5: C.I populations originate from an admixture event between the sampled H.N and an unsampled ghost population; Scenario 6: C.I populations originate from an admixture event between the sampled C.N populations and an unsampled ghost population. As the STRUCTURE results revealed a subsample of the invasive South African *M. dolomieu* individuals to be more closely

related to the historic native samples than to the remaining SA individuals (populations BE and OL; Figure 2.6B), six additional scenarios were simulated (Figure 2.7: A - F). Scenario A: Most of the C.I individuals and the subsample of SA individuals (C.I$_S$) are more closely related to one another than to any other population, but originated from H.N stock which came from the C.N gene pool. Scenario B: Both C.I and C.I$_S$ individuals are closest related to one another, while C.N and H.N are more closely related to one another. Both invasive (C.I and C.I$_S$) and native (C.N and H.N) groupings stem from a communal source population. Scenario C, like Scenario A, states that C.I and C.I$_S$ are most closely related, originating from the C.N population. Both C.N and C.I + C.I$_S$ populations, in turn, originating from the H.N stock. Scenario D proposes a closer tie between H.N and C.I$_S$. This grouping (H.N + C.I$_S$) along with C.I individuals originated from a C.N population. In scenario E, the H.N and C.I$_S$ are once again closest related to one another, originating from C.N. The Remaining C.I individuals along with the H.N + C.I$_S$ + C.N grouping originate from an unsampled population. Lastly scenario F supports the STRUCTURE results, and states that H.N and C.I$_S$ are most closely related, while C.I and C.N are more closely related. Both groupings (H.N + C.I$_S$ and C.I + C.N) share an unsampled origin. To prevent over-parameterization, parameters were specified according to the program guidelines (Cornuet et al. 2014). Firstly, to ensure that at least one scenario and its associated priors could generate simulated data sets similar to that of the observed, a pre-evaluation was performed. This was done by simulating 100, 000 data sets and comparing summary statistics for both single-sample (i.e. mean number of alleles, genetic diversity and allele size variance across loci) and two-sample statistics (i.e. mean genetic diversity, number of alleles, allele size variance, mean index of classification, shared allele distance, distance between samples and F_{ST}) of these simulated data sets to the observed data set (Cornuet et al. 2014). As the mean *M* index across loci (Garza & Williamson 2001) was initially developed with conservation planning in mind, this statistic does not do well with small, unequal sampling sizes and small starting population sizes (Garza & Williamson 2001). Hence, it was excluded from the summary statistics used in the current analyses. Next, I simulated 10^6 data sets per scenario before calculating each scenarios' posterior probability (PP). Scenarios were subsequently compared through a logistic regression, which was conducted on the linear discriminant analysis components (Cornuet et al. 2014). Each scenarios error rate was evaluated by generating 100 pseudo-observed data sets, using parameter values obtained from one of the scenarios (e.g. scenario 1). The type I error rate was determined by counting the number of times the PPs were higher for any scenario other than the chosen scenario, divided by the number of pseudo-observed data sets (i.e. 100), while the type II error rate was calculated

by counting the number of pseudo-observed data sets that unrightfully received the highest PP support (Cornuet et al. 2010).

RESULTS

Contemporary mtDNA analyses

A total of 292 *M. dolomieu* specimens, collected from eight river systems in the invasive SA range were successfully sequenced for 306bp of cytochrome b (cytb) and 979bp of Control Region (CR), resulting in 254 and 857 polymorphic sites, respectively. The nine native USA localities yielded a total of 209 and 174 successfully sequenced *M. dolomieu* specimens for cytb and CR, respectively, with a total of 207 and 730 polymorphic sites retrieved for cytb and CR, respectively. Both cytb and CR rendered fewer haplotypes for the native USA range when compared to the invasive SA range, but similar haplotype and nucleotide diversity levels were observed in both the native and invasive range (Table 2.3). Overall, haplotype and nucleotide diversity levels were high for both native (cytb: $h = 0.976 \pm 0.005$, $\pi = 0.051 \pm 0.025$; CR: $h = 0.977 \pm 0.007$, $\pi = 0.044 \pm 0.021$) and invasive (cytb: $h = 0.967 \pm 0.007$, $\pi = 0.087 \pm 0.043$; CR: $h = 0.985 \pm 0.003$, $\pi = 0.039 \pm 0.019$) range, but differed greatly between sampling localities and gene fragment (native cytb: $h = 0.923$ - 1.000, $\pi = 0.022$ - 0.156; CR: $h = 0.884$ - 1.000, $\pi = 0.001$ - 0.301; invasive cytb: $h = 0.756$ - 0.987, $\pi = 0.033$ - 0.263; CR: $h = 0.867$ - 1.000, $\pi = 0.013$ - 0.088) (Table 2.3). Significant deviations from neutrality were observed for Tajima's *D* and Fu's F_S in both native and invasive range and both gene fragments (Table 2.3). Assessment of population structure using parsimonious haplotype networks showed no distinct geographical clusters in either gene fragment or sampling region (native and invasive range), nor were any haplotypes shared between the native and invasive range (Figure 2.3.1, 2.3.2). Although a large proportion of the specimens in both native and invasive ranges formed part of a star-shaped topology, many mutational steps were observed between other haplotypes (Figure 2.3.1, 2.3.2). Pairwise F_{ST} measures revealed two significantly differentiated groupings, namely invasive SA and native USA (Table 2.4), with comparisons between localities from the two groups ranging from 0.013 to 0.172 for cytb (DO - SAR and KO - VES) and 0.013 to 0.125 for CR (KR - NIA and BE - LOL). Significant within grouping differentiation (though markedly less so for the USA cytb), was also observed in both cytb [(0.014 to 0.150; BR - BU and BE - KO) (0.016 to 0.031; LOL

- NIA / HUD - ONEI and ONEI - VES)] and CR [(0.007 to 0.093; BR - BU and BE - MP) (0.019 to 0.079; HUD - VES and LOL - VES)] (Table 2.4).

Table 2.3. Genetic diversity indices (haplotype (***h***) and nucleotide (***π***)) and neutrality tests (Tajima's neutrality test ***D*** and Fu's neutrality test (***Fs***)) for each of the partial mtDNA gene fragments, namely cytb and CR. Sample size is denoted by **n**, while **H** refers to the number of haplotypes. Statistically significant results ($P < 0.05$) are indicted in bold.

		Cytochrome b (cytb)						**Control Region (CR)**					
		n	**H**	***h***	***π***	***D***	***Fs***	**n**	**H**	***h***	***π***	***D***	***Fs***
Invasive SA Localities	**BE**	20	16	0.963 ± 0.033	0.066 ± 0.034	**-1.682**	-1.758	21	14	0.867 ± 0.074	0.088 ± 0.044	**-2.277**	6.160
	BR	42	33	0.976 ± 0.014	0.061 ± 0.031	-1.295	**-9.88**	43	33	0.981 ± 0.011	0.036 ± 0.018	**-2.011**	-4.340
	BU	47	30	0.965 ± 0.013	0.061 ± 0.031	**-2.004**	-4.574	47	35	0.984 ± 0.008	0.020 ± 0.010	**-2.594**	**-10.918**
	DO	35	30	0.987 ±0.012	0.263 ± 0.129	0.314	-1.295	36	30	0.979 ± 0.016	0.084 ± 0.041	**-2.537**	0.321
	KO	46	24	0.756 ± 0.071	0.044 ± 0.022	**-2.310**	-2.777	45	36	0.984 ± 0.010	0.013 ± 0.007	**-1.71**	**-21.924**
	KR	14	9	0.835 ± 0.101	0.050 ± 0.027	**-1.768**	0.833	15	15	1.000 ± 0.024	0.046 ± 0.024	**-2.047**	-2.642
	MP	45	37	0.987 ± 0.009	0.071 ± 0.036	-0.257	**-11.881**	45	31	0.942 ± 0.024	0.063 ± 0.031	**-2.646**	0.974
	OL	43	24	0.947 ± 0.020	0.033 ± 0.017	**-2.071**	-5.458	40	17	0.906 ± 0.029	0.045 ± 0.022	**-1.603**	8.417
	Overall	292	176	0.967 ± 0.007	0.087 ± 0.043	**-1.899**	**-23.547**	292	179	0.985 ± 0.003	0.039 ± 0.019	**-2.717**	**-23.604**
Native USA Localities	**DET**	7	7	1.000 ± 0.076	0.144 ± 0.083	0.767	-0.226	-	-	-	-	-	-
	HUD	20	15	0.968 ± 0.025	0.050 ± 0.026	**-2.140**	-1.675	17	17	1.000 ± 0.020	0.134 ± 0.068	0.692	-1.145
	LOL	20	16	0.974 ± 0.025	0.040 ± 0.021	**-1.940**	-3.662	20	13	0.884 ± 0.067	0.001 ± 0.001	-1.174	**-15.968**
	NIA	48	31	0.957 ± 0.018	0.032 ± 0.017	**-2.445**	**-12.403**	38	28	0.976 ± 0.014	0.011 ± 0.006	**-2.157**	**-13.583**
	ONEI	30	26	0.989 ± 0.013	0.022 ± 0.012	**-1.545**	**-20.166**	18	17	0.994 ± 0.021	0.082 ± 0.042	**-2.389**	-0.867
	ONEO	10	8	0.956 ± 0.059	0.156 ± 0.084	-0.689	2.782	10	10	1.000 ± 0.045	0.012 ± 0.007	**-1.575**	**-4.188**
	SAR	13	12	0.987 ± 0.035	0.030 ± 0.017	-0.615	**-4.471**	7	7	1.000 ± 0.076	0.301 ± 0.169	**-1.806**	2.179
	STL	47	34	0.966 ± 0.017	0.032 ± 0.017	-0.829	**-18.178**	51	32	0.942 ± 0.023	0.002 ± 0.001	**-1.960**	**-28.464**
	VES	14	10	0.923 ± 0.060	0.022 ± 0.012	**-1.950**	-2.114	13	10	0.962 ± 0.041	0.059 ± 0.031	-1.418	2.703
	Overall	209	116	0.976 ± 0.005	0.051 ± 0.025	**-2.191**	**-23.870**	174	117	0.977 ± 0.007	0.044 ± 0.021	**-1.829**	**-23.756**

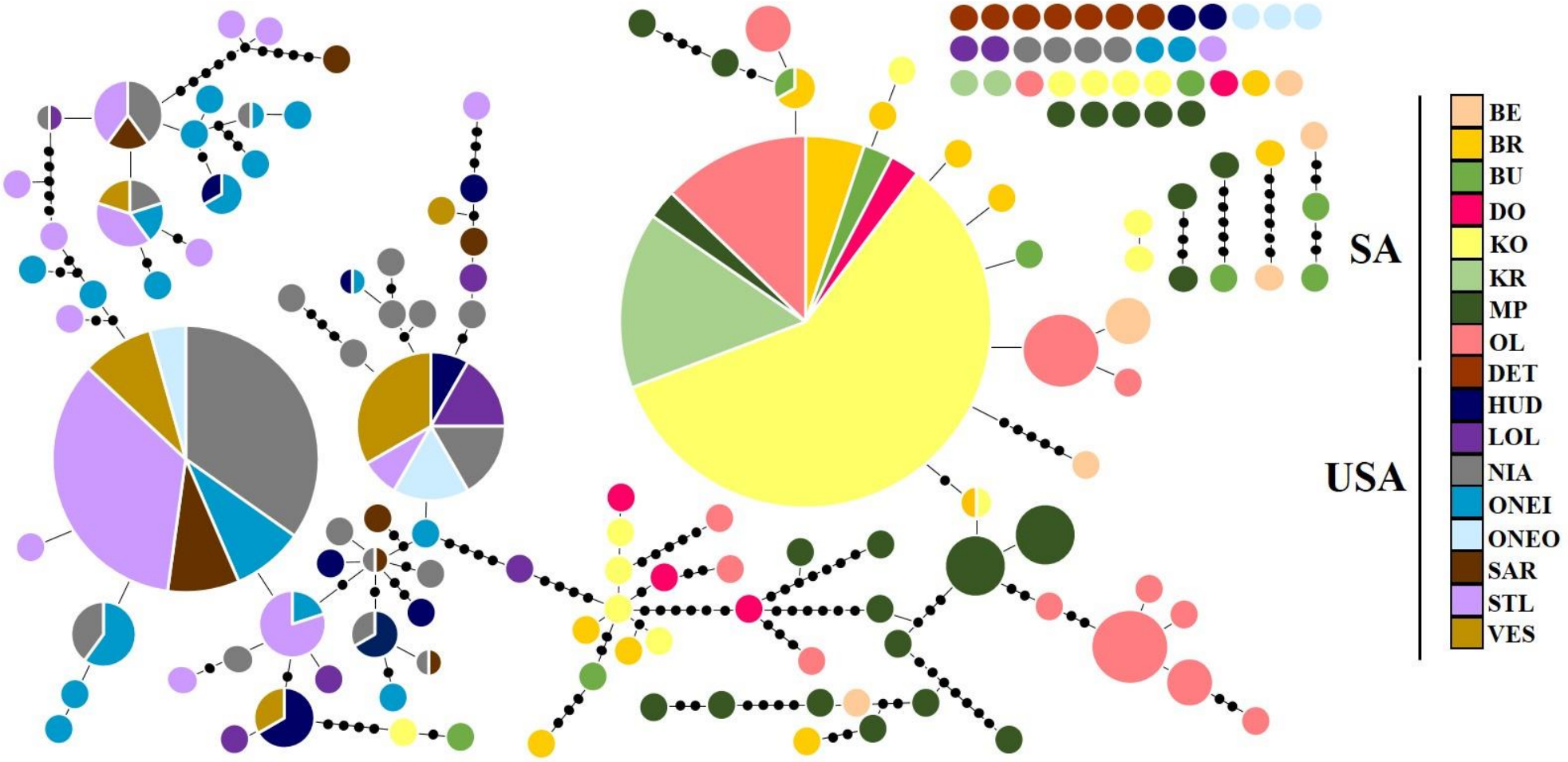


Figure 2.3.1. A haplotype network for cytb constructed with TCS using parsimony and based on the combined contemporary native USA and invasive SA mtDNA dataset. The size of each haplotype is proportional to its frequency, while each mutational step, missing haplotype or unsampled haplotype is represented by a black dot. Colours correspond to the sampled localities, as represented in the legend.

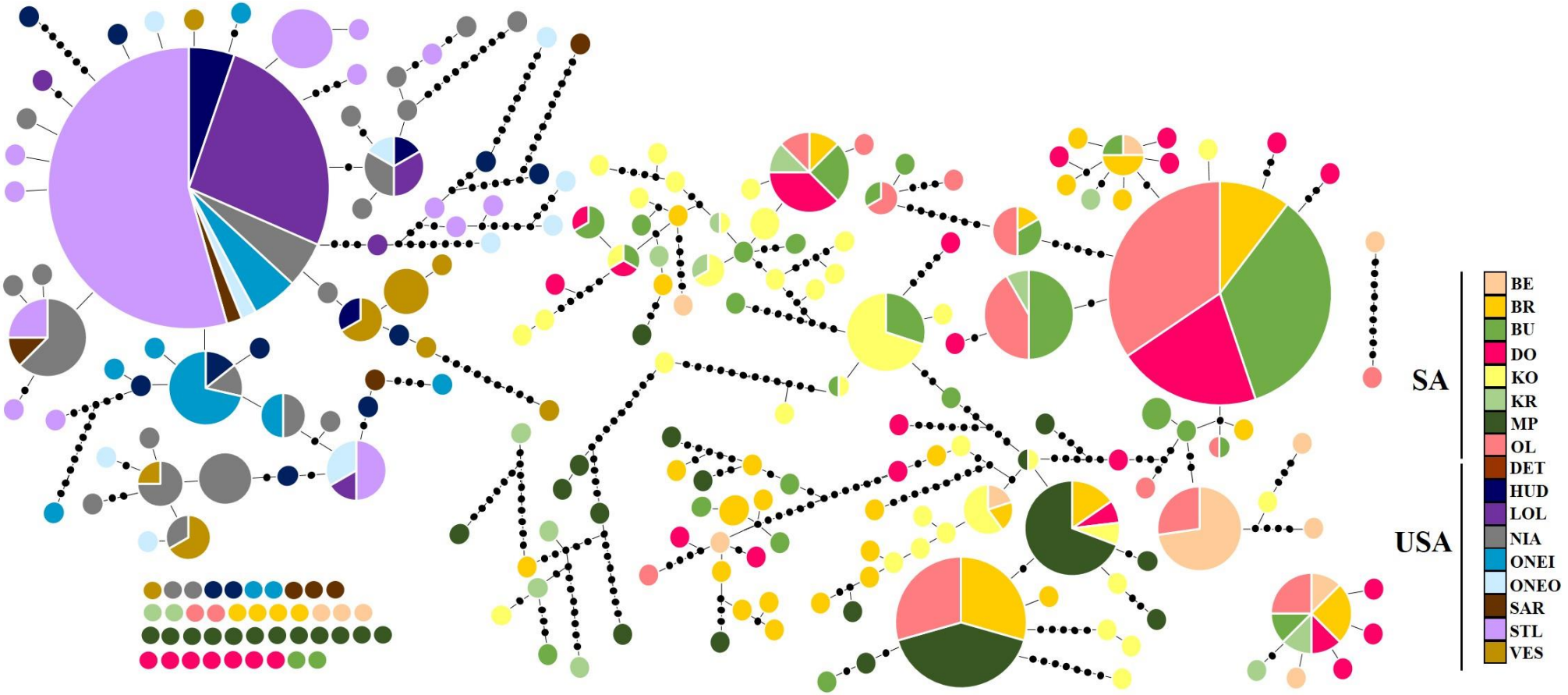


Figure 2.3.2. A parsimonious haplotype network constructed in TCS for CR based on the combined contemporary native USA and invasive SA mtDNA dataset. Each haplotypes size is proportional to its frequency, and each mutational step, unsampled haplotype or missing haplotype is represented by a black dot. Colours correspond to the sampled localities, as represented in the legend.

Table 2.4. Pairwise F_{ST} values between native and invasive *M. dolomieu* populations based on the two partial mtDNA gene fragments, cytb (below diagonal) and CR (above diagonal). Statistically significant results in bold ($P < 0.05$).

	BE	**BR**	**BU**	**DO**	**KO**	**KR**	**MP**	**OL**	**DET**	**HUD**	**LOL**	**NIA**	**ONEI**	**ONEO**	**SAR**	**STL**	**VES**
BE	0.000	**0.070**	**0.070**	**0.075**	**0.068**	**0.069**	**0.093**	**0.086**	-	**0.068**	**0.125**	**0.076**	**0.074**	**0.072**	0.075	**0.093**	**0.088**
BR	**0.033**	0.000	**0.007**	0.006	**0.015**	0.005	**0.011**	**0.020**	-	0.010	**0.065**	**0.021**	**0.016**	0.010	0.011	**0.039**	**0.028**
BU	**0.024**	**0.014**	0.000	0.002	**0.015**	0.000	**0.037**	**0.020**	-	0.008	**0.063**	**0.020**	**0.014**	0.009	0.009	**0.037**	**0.026**
DO	**0.016**	**0.016**	**0.020**	0.000	**0.018**	0.003	**0.034**	**0.020**	-	0.011	**0.066**	**0.022**	**0.017**	0.011	0.012	**0.040**	**0.029**
KO	**0.150**	**0.080**	**0.091**	**0.116**	0.000	0.004	**0.032**	**0.055**	-	0.008	**0.063**	**0.020**	**0.015**	0.009	0.009	**0.037**	**0.027**
KR	**0.101**	**0.041**	**0.052**	**0.072**	-0.010	0.000	**0.031**	**0.037**	-	0.000	**0.059**	**0.013**	0.007	0.000	0.000	**0.031**	**0.019**
MP	**0.025**	0.005	**0.017**	**0.015**	**0.121**	**0.076**	0.000	**0.057**	-	**0.030**	**0.085**	**0.041**	**0.036**	**0.032**	0.033	**0.058**	**0.049**
OL	**0.034**	0.009	**0.020**	**0.016**	**0.093**	**0.053**	**0.015**	0.000	-	**0.049**	**0.104**	**0.059**	**0.055**	**0.052**	**0.054**	**0.076**	**0.068**
DET	0.023	0.014	0.020	0.008	**0.149**	0.091	0.010	0.015	0.000	-	-	-	-	-	-	-	-
HUD	**0.037**	**0.028**	**0.033**	**0.022**	**0.147**	**0.096**	**0.024**	**0.029**	0.017	0.000	**0.036**	0.005	-0.003	0.000	-0.008	0.011	**0.019**
LOL	**0.034**	**0.025**	**0.031**	**0.020**	**0.144**	**0.093**	**0.022**	**0.026**	0.015	-0.006	0.000	**0.042**	**0.038**	**0.062**	0.015	0.001	**0.079**
NIA	**0.042**	**0.033**	**0.038**	**0.028**	**0.143**	**0.097**	**0.029**	**0.034**	0.024	**0.025**	**0.016**	0.000	0.007	0.013	-0.017	**0.020**	**0.031**
ONEI	**0.026**	**0.018**	**0.024**	**0.012**	**0.132**	**0.083**	**0.014**	**0.019**	0.006	**0.016**	**0.019**	0.010	0.000	0.007	-0.001	0.016	**0.025**
ONEO	**0.043**	**0.033**	**0.039**	**0.028**	**0.162**	**0.107**	**0.029**	**0.035**	0.023	-0.013	-0.032	0.008	**0.020**	0.000	0.000	**0.030**	0.020
SAR	**0.028**	**0.019**	**0.025**	**0.013**	**0.144**	**0.090**	**0.015**	**0.020**	0.007	0.011	0.008	-0.008	0.002	-0.003	0.000	-0.004	0.021
STL	**0.038**	**0.029**	**0.035**	**0.024**	**0.139**	**0.093**	**0.026**	**0.030**	0.019	**0.027**	**0.021**	0.002	0.008	0.009	-0.007	0.000	**0.049**
VES	**0.059**	**0.049**	**0.054**	**0.043**	**0.172**	**0.121**	**0.045**	**0.050**	0.042	0.019	-0.003	0.010	**0.031**	-0.035	0.018	0.017	0.000

The AMOVA, conducted independently on each mtDNA gene fragment, revealed that the largest proportion of genetic variation (cytb: 94.79 %; CR: 95.79 %) was distributed within each population, with very little variation observed between the invasive SA and native USA groups (cytb: 2.15 %; CR: 1.58 %) as well as among populations within groups (cytb: 3.06 %; CR: 2.26 %). All variance components were, however, significant ($P < 0.001$) (Table 2.5).

Table 2.5. Analysis of molecular variance (AMOVA) results performed on native USA and invasive SA populations of *M. dolomieu* for both partial mtDNA gene regions.

	Source of Variation	**d.f.**	**Sum of Squares**	**Variance Components**	**% Variation**	***P* value**
Cytb	Among groups	1	3.639	**0.011**	2.15	<0.001
	Among populations within groups	15	13.66	**0.015**	3.06	<0.001
	Within populations	484	229.296	**0.474**	94.79	<0.001
	Total	500	246.595	0.500		
CR	Among groups	1	2.659	**0.008**	1.58	<0.001
	Among populations within groups	14	11.9	**0.013**	2.26	<0.001
	Within populations	450	215.761	**0.479**	95.79	<0.001
	Total	465	230.32	0.500		

Contemporary and historic microsatellite analyses

A total of 519 contemporary sampled specimens, representing both the invasive SA (n = 306; eight localities) and native USA (n = 213; nine localities) ranges, were successfully genotyped for all nine microsatellite loci, while 53 museum samples, representing 11 localities within the historical native range, were successfully genotyped for eight microsatellite loci. Neither of the three groups (contemporary invasive SA – C.I, contemporary native USA – C.N, historic native USA – H.N) displayed any amplification errors (i.e. large allele dropout, stuttering), nor did any loci exhibit linkage disequilibrium. FreeNA (Chapuis & Estoup 2007), using the EM algorithm (Dempster et al. 1977), revealed the presence of null alleles in microsatellite Mdo9 within the historic native USA samples, but this was not the case for either of the contemporary groups. In addition, most of the populations, groups and loci exhibited significant deviations from Hardy – Weinberg expectations (Table 2.6). The number of alleles (Na) was comparable between localities in both C.N and C.I datasets, while the H.N dataset rendered a two-to-three-fold increase in Na for the majority of loci, with similar results being observed for allelic richness (AR) following the rarefaction analysis (Table 2.6). Multi-locus genetic diversity (observed heterozygosity, H_O) ranged from 0.39 (ONEI) to 0.59 (DET), while levels of expected heterozygosity (H_E) ranged from 0.35 (MP) to 0.73 (MUSEUM) across all loci. Significant levels ($P < 0.05$) of inbreeding (F_{IS}) were observed for two C.I populations (BE = 0.26, OL = 0.17), as well as the H.N population (MUSEUM = 0.43), with no significant levels of inbreeding being detected in the C.N populations.

A stacked bar graph (Figure 2.4) revealed substantial variation in observed heterozygosity (H_O) among populations and loci, with reservoirs (such as dams and lakes with a catchment size < 5000 km^2) consistently displaying lower levels of H_O. The ANOVA revealed significant differences between the thee groups ($F_{2,214} = 22.90$, $P = < 0.001$), with H_O being higher in the historic native group compared to both contemporary groups (Bonferroni *post hoc* test $P < 0.001$). However, a significant marker effect ($F_{7,214} = 19.82$, $P < 0.001$) was observed. Pairwise F_{ST} values revealed significant population differentiation among C.I populations, ranging from 0.066 – 0.469 (DO – KO and BE - MP), with similar results being observed when comparing populations across all three groups, i.e. C.I, C.N and H.N (0.123 – 0.537; MP – SAR and OL – MUSEUM) (Table 2.7). In contrast, the C.N populations displayed significantly less population differentiation among sampled localities within this group (0.072 – 0.129; LOL – NIA and SAR – STL) (Table 2.7).

The Principal Component Analysis (PCA), based on allelic frequencies, revealed two distinct groups along the first two axes; the first comprising both C.N and C.I populations and the second comprising H.N populations (Figure 2.5). Limited genetic associations between the two groups were observed. The Bayesian clustering analyses conducted in STRUCTURE revealed extensive population sub-structuring within the C.I localities (BE – OL), with Delta K (Evanno et al. 2005) retrieving $K = 5$ as the most probable number of clusters (Figure 2.6A). Both reservoirs (BU and MP) were represented by their own cluster and showed very little population variation, corroborating the genetic diversity results (Figure 2.4; Table 2.6). The remaining six C.I populations, however, displayed substantial levels of admixture, in particular localities BE and OL (Figure 2.6A). The C.N populations (DET – VES) exhibited high levels of population admixture indicative of shallow population differentiation, with Delta K revealing the most probable $K = 4$ (Figure 2.6A). Similar levels of admixture and Delta K ($K = 4$) were obtained for the H.N populations (PO – FC) (Figure 2.6A). To determine the most probable source population of the C.I populations, all 28 localities representing all three groups (C.I, C.N, H.N), were combined (Figure 2.6B). Delta K revealed the most probable number of clusters to be $K = 3$, with each cluster representing a group, though admixture between the two contemporary groups was observed. Interestingly, a subset of individuals within the BE and OL (and to a lesser extent DO and KO) contemporary invasive localities shared a cluster with the H.N group, but this was not the case for any of the C.N populations, despite overlapping sampling localities (DET, HUD, Susquehanna River: LOL, ONEO, VES, SU; Table 2.1) (Figure 2.6B).

Table 2.6. Genetic diversity measures for all 18 localities (contemporary invasive SA = BE – OL, historic native USA = MUS, contemporary native USA = DET – VES) at nine microsatellite loci: **n** – number of successfully genotyped individuals; **Na** – number of alleles; **AR** – allelic richness following the rarefaction analysis; $\mathbf{H_E}$ – expected heterozygosity; $\mathbf{H_O}$ – observed heterozygosity; $\mathbf{F_{IS}}$ – inbreeding coefficient. Statistical significant results in bold ($P < 0.05$).

			LOCALITY																	
			BE	**BR**	**BU**	**DO**	**KO**	**KR**	**MP**	**OL**	**MUS**	**DET**	**HUD**	**LOL**	**NIA**	**ONEI**	**ONEO**	**SAR**	**STL**	**VES**
Mdo3	571	**n**	22	43	48	38	46	15	50	44	52	7	21	20	49	27	10	10	55	14
	10	**Na**	3	3	1	4	4	2	1	3	10	3	3	3	3	3	3	3	3	3
	2.73	**AR**	2.86	2.55	1.00	2.52	3.03	1.94	1.00	2.78	3.92	2.95	2.82	2.40	2.72	2.28	2.16	2.92	2.57	2.69
	0.48	**H_E**	0.62	0.58	0.00	0.48	0.64	0.37	0	0.61	0.73	0.63	0.62	0.38	0.55	0.48	0.28	0.65	0.52	0.58
	0.41	**H_O**	0.27	0.60	0.00	0.39	0.65	0.33	0	0.39	0.46	0.71	0.62	0.25	0.51	0.33	0.30	0.60	0.49	0.43
	0.17	**F_{IS}**	**0.57**	-0.05	\	0.18	-0.02	0.10	\	**0.37**	**0.37**	-0.15	-0.01	0.35	0.07	0.30	-0.08	0.08	0.06	0.27
Mdo4	565	**n**	22	43	48	35	46	15	50	44	51	7	20	20	49	26	10	10	55	14
	17	**Na**	2	2	2	3	3	2	1	3	16	3	2	2	4	2	2	2	2	2
	2.30	**AR**	1.99	1.76	1.51	1.76	2.00	1.98	1.00	2.71	4.61	2.53	2.00	2.00	2.27	2.00	2.00	1.99	1.97	1.99
	0.41	**H_E**	0.49	0.24	0.14	0.21	0.31	0.43	0.00	0.61	0.79	0.38	0.51	0.51	0.43	0.51	0.51	0.44	0.44	0.45
	0.38	**H_O**	0.23	0.28	0.15	0.17	0.28	0.47	0.00	0.45	0.37	0.43	0.50	0.70	0.43	0.42	0.80	0.40	0.60	0.21
	0.11	**F_{IS}**	**0.54**	-0.15	-0.07	0.18	0.09	-0.08	\	**0.25**	**0.53**	-0.13	0.02	-0.38	0.01	0.17	-0.64	0.10	-0.38	0.54
Mdo5	558	**n**	22	42	48	38	46	15	49	44	44	7	21	20	48	27	10	10	53	14
	20	**Na**	3	3	2	5	4	2	4	5	10	1	5	3	4	3	3	3	4	5
	2.83	**AR**	2.63	2.20	1.93	2.74	2.13	1.91	2.56	3.46	4.17	1.00	2.78	2.40	2.92	1.33	2.42	2.57	2.54	3.23
	0.45	**H_E**	0.51	0.52	0.38	0.59	0.41	0.33	0.57	0.68	0.75	0.00	0.44	0.54	0.55	0.07	0.47	0.43	0.39	0.56
	0.40	**H_O**	0.27	0.36	0.50	0.63	0.30	0.27	0.69	0.55	0.30	0.00	0.52	0.65	0.48	0.07	0.60	0.30	0.32	0.36
	0.14	**F_{IS}**	**0.47**	**0.31**	-0.32	-0.07	**0.26**	0.20	-0.23	**0.20**	**0.61**	\	-0.19	-0.21	0.13	-0.01	-0.30	0.31	0.17	**0.37**

Table 2.6 continued on next page

Table 2.6 continued

Mdo7	571	**n**	22	43	48	38	46	15	50	44	52	7	21	20	49	27	10	10	55	14
	11	**Na**	3	4	4	6	4	4	3	4	11	3	5	3	5	3	3	2	3	3
	3.26	**AR**	2.59	3.11	3.49	3.64	3.70	3.52	2.48	3.59	4.52	2.86	2.80	2.22	2.98	2.16	2.45	1.71	2.08	2.79
	0.58	**H_E**	0.47	0.65	0.69	0.7	0.75	0.72	0.43	0.73	0.79	0.56	0.57	0.51	0.59	0.51	0.54	0.19	0.51	0.59
	0.52	**H_O**	0.41	0.51	0.73	0.71	0.72	0.80	0.46	0.64	0.46	0.29	0.67	0.55	0.51	0.33	0.50	0.00	0.58	0.50
	0.10	**F_{IS}**	0.13	**0.21**	-0.05	-0.01	0.04	-0.12	-0.06	0.12	**0.42**	0.51	-0.17	-0.08	0.13	**0.36**	0.08	1.00	-0.14	0.16
Mdo8	534	**n**	22	43	48	37	46	15	50	44	-	7	21	20	49	25	10	9	55	14
	17	**Na**	6	4	2	6	4	3	2	7	-	5	5	5	6	5	4	3	3	6
	3.78	**AR**	4.42	3.62	1.89	3.72	2.75	2.30	1.49	3.46	-	4.15	3.32	3.01	4.07	2.87	3.13	2.49	2.78	3.44
	0.58	**H_E**	0.80	0.73	0.33	0.72	0.54	0.35	0.13	0.63	-	0.73	0.58	0.59	0.76	0.50	0.51	0.50	0.62	0.53
	0.53	**H_O**	0.91	0.70	0.38	0.65	0.43	0.33	0.10	0.52	-	0.71	0.38	0.60	0.76	0.48	0.40	0.44	0.71	0.50
	0.07	**F_{IS}**	-0.14	0.05	-0.13	0.10	0.20	0.05	0.24	**0.17**	-	0.02	0.35	-0.02	0.01	0.05	0.22	0.12	-0.14	0.07
Mdo9	569	**n**	22	43	48	38	46	15	50	43	51	7	21	20	49	27	10	10	55	14
	11	**Na**	6	3	3	4	4	3	2	4	10	2	2	2	3	2	2	2	3	2
	2.60	**AR**	3.51	2.68	2.24	2.27	2.37	2.87	1.78	3.08	4.27	1.97	2.00	1.90	2.09	2.00	1.97	1.97	2.08	2.00
	0.49	**H_E**	0.64	0.60	0.49	0.42	0.37	0.63	0.26	0.63	0.76	0.36	0.49	0.33	0.51	0.51	0.39	0.39	0.51	0.49
	0.46	**H_O**	0.59	0.63	0.60	0.39	0.35	0.60	0.18	0.58	0.27	0.14	0.62	0.40	0.47	0.41	0.50	0.50	0.51	0.50
	0.10	**F_{IS}**	0.08	-0.04	-0.24	0.07	0.06	0.04	0.30	0.07	**0.64**	0.63	-0.26	-0.23	0.09	0.20	-0.29	-0.29	0.01	-0.01
Mdo10	572	**n**	22	43	48	38	46	15	50	44	53	7	21	20	49	27	10	10	55	14
	6	**Na**	2	2	2	2	2	2	2	2	6	1	2	1	2	2	1	1	2	2
	1.71	**AR**	1.37	1.79	1.40	1.22	1.70	1.99	1.98	1.89	2.27	1.00	1.94	1.00	1.74	1.61	1.00	1.00	1.83	2.00
	0.22	**H_E**	0.09	0.26	0.10	0.05	0.21	0.48	0.45	0.33	0.33	0.00	0.37	0.00	0.23	0.17	0.00	0.00	0.29	0.52
	0.22	**H_O**	0.09	0.30	0.10	0.05	0.24	0.33	0.46	0.27	0.23	0.00	0.29	0.00	0.27	0.19	0.00	0.00	0.31	0.86
	0.00	**F_{IS}**	-0.02	-0.17	-0.04	-0.01	-0.13	0.31	-0.03	0.17	**0.32**	\	0.24	\	-0.14	-0.08	\	\	-0.07	-0.70

Table 2.6 continued on next page

Table 2.6 continued

Mdo11	571	**n**	22	43	48	38	46	15	50	44	52	7	21	20	49	27	10	10	55	14
	18	**Na**	3	2	2	5	4	2	3	3	9	4	2	3	3	3	2	2	2	2
	2.54	**AR**	2.36	1.98	1.72	2.92	2.87	1.98	1.35	2.82	4.36	3.29	1.94	1.97	2.31	2.10	1.93	2.00	1.86	2.00
	0.45	**H_E**	0.54	0.44	0.22	0.50	0.63	0.43	0.08	0.61	0.80	0.66	0.37	0.27	0.53	0.40	0.34	0.52	0.31	0.52
	0.40	**H_O**	0.23	0.37	0.25	0.50	0.59	0.60	0.08	0.32	0.46	0.86	0.29	0.20	0.57	0.30	0.40	0.30	0.35	0.50
	0.16	**F_{IS}**	**0.58**	0.16	-0.13	0.00	0.07	-0.40	-0.02	**0.48**	**0.43**	-0.33	0.24	0.26	-0.09	0.26	-0.20	0.44	-0.11	0.03
Lma21	564	**n**	22	42	48	38	46	15	49	44	48	7	21	20	49	27	10	10	54	14
	24	**Na**	7	6	4	8	6	4	3	7	22	4	4	3	5	5	3	5	6	5
	4.15	**AR**	4.47	3.73	2.50	4.96	3.31	3.15	2.44	4.56	5.85	3.94	3.23	2.73	3.08	3.36	2.45	3.80	3.40	3.73
	0.69	**H_E**	0.77	0.68	0.46	0.83	0.66	0.64	0.56	0.79	0.90	0.80	0.66	0.61	0.67	0.69	0.57	0.73	0.69	0.73
	0.88	**H_O**	0.68	0.81	0.52	0.95	0.87	0.87	0.90	0.93	0.79	1.00	1.00	1.00	1.00	0.96	1.00	0.80	0.89	0.86
	-0.26	**F_{IS}**	0.12	-0.19	-0.13	-0.15	-0.32	-0.37	-0.61	-0.18	**0.12**	-0.27	-0.54	-0.67	-0.50	-0.40	-0.82	-0.11	-0.28	-0.18
Ave. all loci		**n**	22	43	48	38	46	15	50	44	50	7	21	20	49	27	10	10	55	14
		Na	4	3	3	5	4	3	3	4	12	3	3	3	4	3	3	3	3	3
		AR	2.91	2.60	1.96	2.86	2.65	2.40	1.79	3.15	4.25	2.63	2.54	2.18	2.69	2.19	2.17	2.27	2.35	2.65
		H_E	0.55	0.52	0.35	0.50	0.50	0.49	0.35	0.62	0.73	0.59	0.51	0.47	0.54	0.43	0.45	0.48	0.48	0.55
		H_O	0.41	0.51	0.40	0.49	0.49	0.51	0.41	0.52	0.42	0.59	0.54	0.54	0.55	0.39	0.56	0.42	0.53	0.52
		F_{IS}	**0.26**	0.03	-0.15	0.01	0.02	-0.05	-0.16	**0.17**	**0.43**	-0.01	-0.06	-0.17	-0.04	0.09	-0.27	0.14	-0.11	0.06

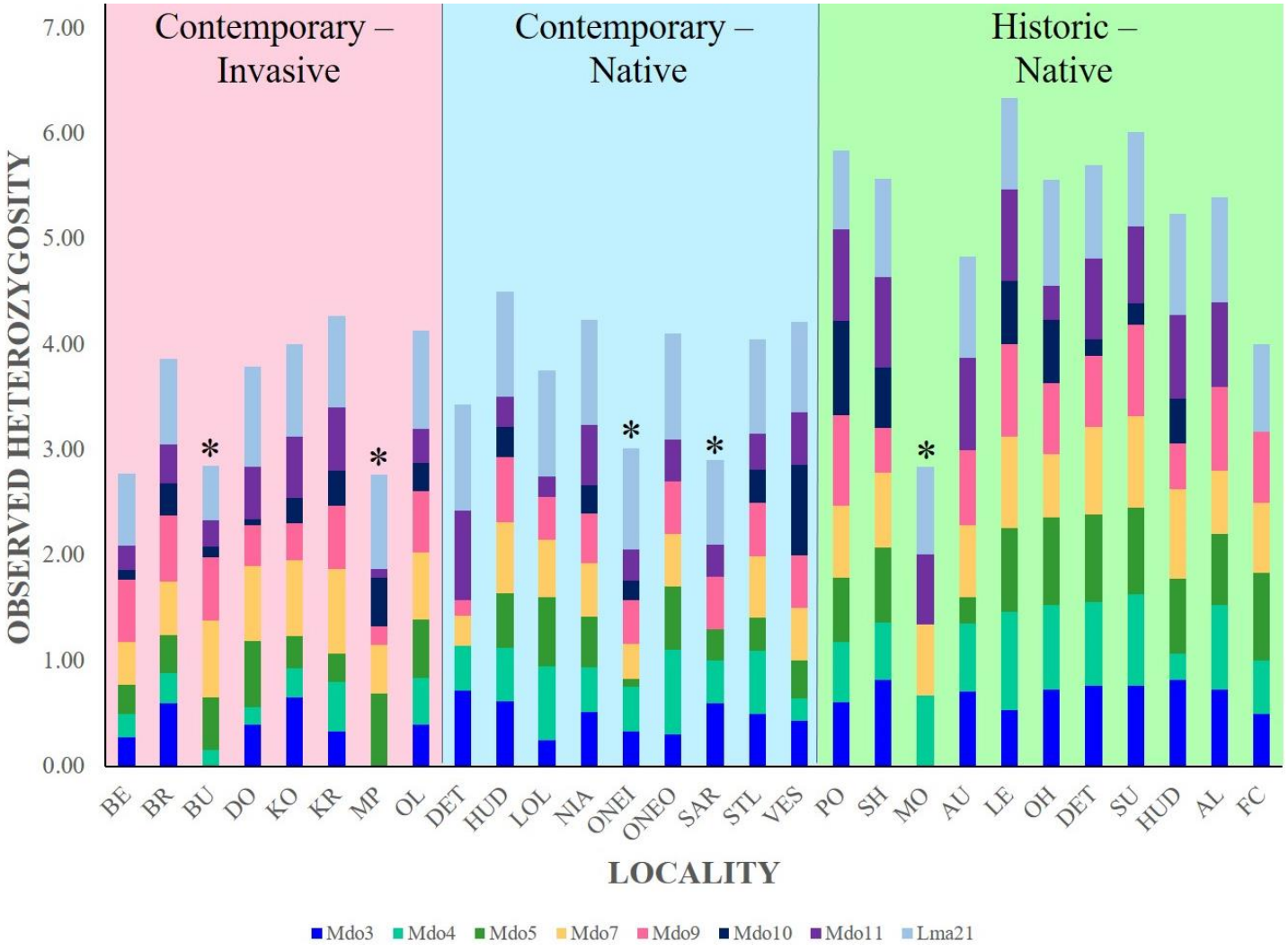


Figure 2.4. A stacked bar graph representing the variation in observed heterozygosity (H_O) among populations and loci between the three groups (contemporary invasive SA, contemporary native USA, historic native USA). Reservoirs (excluding Lake Erie (LE)) are indicated with an asterisks (*).

Table 2.7. Pairwise F_{ST} values between contemporary invasive, contemporary native and historic native *M. dolomieu* populations. Values are based on eight microsatellite loci and statistical significant results are indicated in bold ($P < 0.05$).

	BE	**BR**	**BU**	**DO**	**KO**	**KR**	**MP**	**OL**	**DET**	**HUD**	**LOL**	**NIA**	**ONEI**	**ONEO**	**SAR**	**STL**	**VES**
BE	—																
BR	**0.231**	—															
BU	**0.406**	**0.269**	—														
DO	**0.192**	0.080	**0.211**	—													
KO	**0.229**	**0.088**	**0.237**	**0.066**	—												
KR	**0.274**	**0.108**	**0.389**	**0.139**	0.082	—											
MP	**0.469**	**0.325**	**0.349**	**0.260**	**0.225**	0.309	—										
OL	**0.146**	0.064	**0.269**	**0.101**	**0.071**	**0.097**	**0.274**	—									
DET	**0.276**	0.198	**0.438**	0.247	0.261	0.290	**0.532**	0.210	—								
HUD	**0.278**	0.180	**0.397**	**0.233**	**0.226**	**0.278**	**0.461**	0.159	**0.094**	—							
LOL	**0.331**	**0.209**	**0.429**	**0.265**	**0.261**	**0.339**	**0.527**	**0.205**	0.253	0.139	—						
NIA	**0.239**	0.111	**0.277**	**0.133**	0.119	**0.195**	**0.340**	**0.126**	0.151	0.088	**0.072**	—					
ONEI	**0.306**	0.175	**0.407**	**0.212**	**0.182**	**0.248**	**0.464**	0.167	0.263	0.148	0.073	0.049	—				
ONEO	**0.335**	**0.212**	**0.431**	**0.261**	**0.251**	**0.334**	**0.534**	**0.196**	0.283	0.144	0.004	0.062	0.038	—			
SAR	**0.289**	**0.197**	**0.468**	**0.263**	**0.227**	**0.259**	**0.537**	**0.182**	0.185	0.161	**0.128**	0.093	0.094	0.121	—		
STL	**0.300**	0.129	**0.367**	0.184	0.157	**0.219**	**0.402**	**0.157**	0.218	0.122	**0.102**	0.041	0.037	**0.098**	**0.129**	—	
VES	**0.244**	0.140	**0.336**	0.163	0.158	0.184	**0.374**	0.130	0.188	0.087	0.123	0.040	0.065	0.103	0.131	0.086	—
MUS	**0.171**	**0.170**	**0.315**	**0.180**	**0.191**	**0.224**	**0.382**	**0.123**	0.158	0.160	**0.185**	**0.148**	0.191	**0.189**	**0.168**	**0.196**	**0.143**

The DIYABC analysis consistently supported the notion of a more recent introduction. The first set of scenarios tested (Scenarios 1 – 6; Figure 2.7), revealed that Scenario 2 had the highest posterior probability (direct approach: 0.494, 95% CI 0.558 – 0.932; logistic approach: 0.946, 95% CI 0.930 – 0.962). The second set of analyses (Scenario A - F; Figure 2.7) supported both Scenario C and F, with marginally higher values observed for Scenario F (direct approach: 0.526, 95% CI 0.088 – 0.964; logistic approach: 0.431, 95% CI 0.342 – 0.520). Type I and Type II error rates were marginally low for both scenarios (Scenario 2: Type I error = 0.433; and Type II error = 0.134 (0.060 – 0.252); Scenario F: Type I error = 0.268; and Type II error = 0.071 (0.042 – 0.103)).

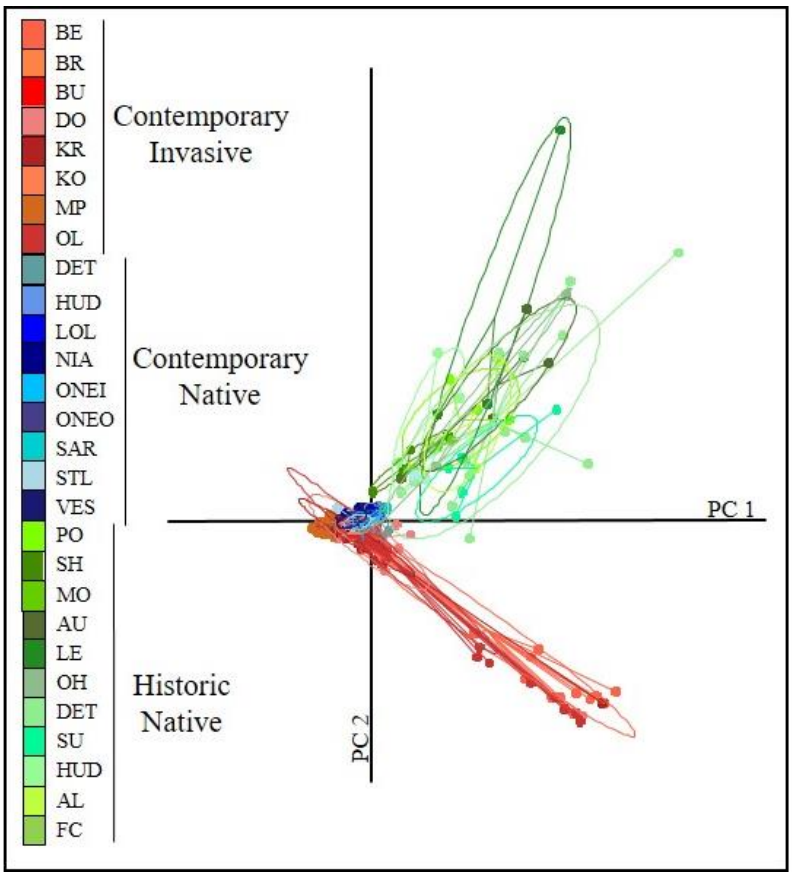


Figure 2.5. PCA analyses conducted on the combined microsatellite genotypes for the three groups (i.e. contemporary invasive, contemporary native, historic native). Each dot represents a genotyped individual, with colours corresponding to sampled localities.

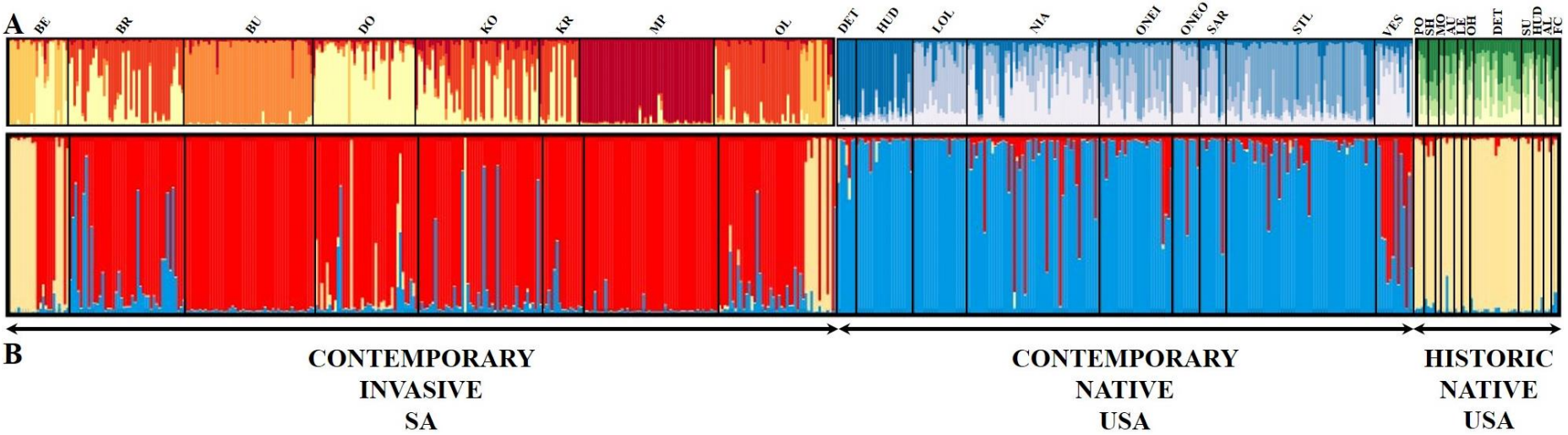


Figure 2.6. STRUCTURE plots representing the population structure within (A) each of the three groups (contemporary invasive SA, contemporary native USA, historic native USA) when run independently, and (B) population structure for all localities combined into a single run. Each genotyped individual is represented by a vertical line, with each lines' colour being proportional to the cluster membership of the individual.

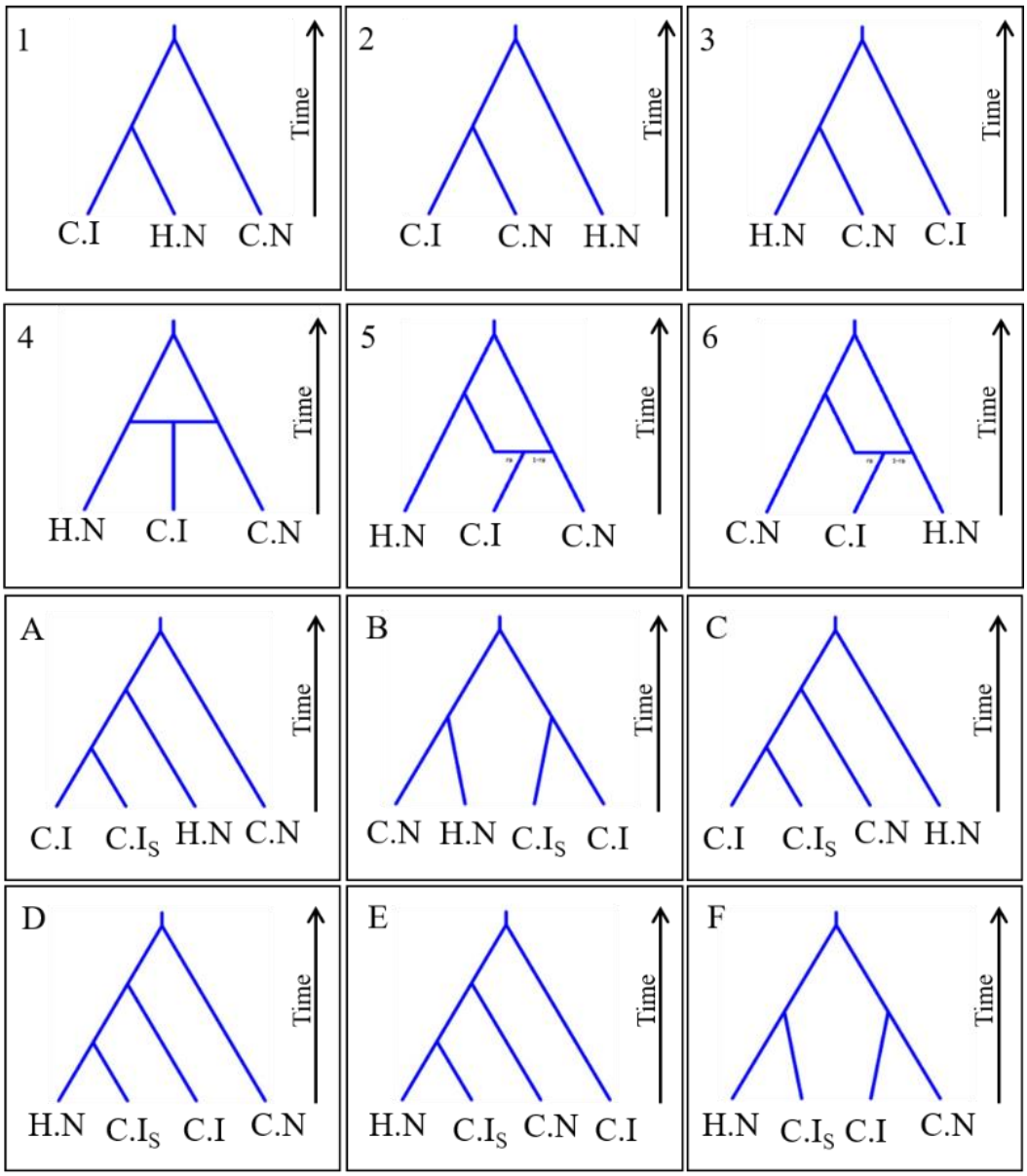


Figure 2.7. Probable introduction scenarios tested with Approximate Bayesian Computation as implemented in DIYABC. C.I – contemporary invasive SA, $C.I_S$ – contemporary invasive SA sub-population, C.N – contemporary native USA, H.N – historic native USA.

DISCUSSION

Understanding the demographic history of an invasive species is thought to be a fundamental constituent of invasion biology. Indeed, numerous studies have compared genetic diversity levels across native and invasive ranges in an attempt to reconstruct the invasion history of invasive species (reviewed in Lee et al. 2004; Novak & Mack 2005; Roman & Darling 2007; Dlugosh & Parker 2008; Rius & Darling 2014). Theory predicts that invasive species would experience founding events (i.e. genetic bottleneck) upon introduction, ultimately leading to a reduction in genetic variation (Nei et al. 1975). Using the smallmouth bass, *Micropterus dolomieu*, as a model system, I predicted that the contemporary invasive (C.I) South African range would have less genetic diversity than the native (historic – H.N and contemporary – C.N) North American range. On the contrary, my results revealed marginally elevated levels of genetic diversity in the C.I range when compared to the C.N range (Table 2.3, 2.6). However, this was not the case when comparing both C.N and C.I groups to the H.N group, as the H.N range displayed the highest levels of heterozygosity, number of alleles (Na) and allelic richness (AR) values (Table 2.6). Although a rarefaction analysis (on the AR), as suggested by Dlugosch & Parker (2008), was used to circumvent the sample size discrepancy, differences in sample size (historical native USA: n = 53, contemporary native USA: n = 213, contemporary invasive SA: n = 306) may still have influenced the results. Similar results were, however, obtained by Nielsen et al. (1997) who observed a significant decrease in the number of alleles (Na) in a contemporary population of Danish Atlantic salmon (*Salmo salar*) when compared to 60-year-old salmon scales from the same population. They concluded that this decrease was likely due a genetic bottleneck in the contemporary population. My results support this hypothesis, as Tajima's *D* and Fu's *Fs* revealed significantly negative values for the C.N populations, in addition to displaying high haplotype- and low nucleotide diversity levels (Table 2.3), indicative of a population that underwent a genetic bottleneck before experiencing a rapid population expansion (Grant & Bowen 1998). Moreover, the lack of population structure in the C.N population (Figure 2.5; 2.6) and the limited AR and Na (Table 2.6) within the microsatellite dataset further support this notion. A possible explanation for the temporal decline in AR and heterozygosity within the contemporary native range may be explained by overfishing and subsequent stocking events within the USA (i.e. changes in population size over time). Early in the 1800's black bass, particularly *Micropterus salmoides* and *M. dolomieu*, were being harvested commercially and recreationally at alarming rates for consumption purposes (Long et al. 2015). As noted by Allendorf et al. (2008), harvesting often

targets specific size classes (which in turn, is closely related to age and sexual maturity), thereby reducing the effective breeding population size (Ne) and increasing the rate at which genetic variation is lost. By the 1870's, the overfishing, in addition to industrial pollution (Marsh 1867), had dramatically decreased the number of black bass populations, prompting the US government to start breeding black bass in hatcheries and enforce stricter policies on fishing (Long et al. 2015). In 1903 alone, 528, 365 black bass were released into depauperated waterbodies across the USA (Bowers 1905; Loppnow et al. 2013; Long et al. 2015). Hence, one can assume that these intense stocking practices could have led to the high levels of admixture observed within C.N *M. dolomieu* individuals and populations, with similar results being observed in brook charr (*Salvelinus fontinalis*) (Lamaze et al. 2012).

Elevated levels of genetic diversity are, however, not uncommon in invasive species in a novel invaded range, and are often attributed to multiple introductions and/ or admixture (see Rius & Darling 2014 for a comprehensive review) as observed in an array of studies (Kolbe et al. 2004; Yonekura et al. 2007; Gillis et al. 2009; Pairon et al. 2010; Funk et al. 2011; Beneteau et al. 2012; Gray et al. 2014; Lippens et al. 2017). The results from the STRUCTURE analyses contradict the theory that invasive South African *M. dolomieu* populations originate from a single introductory event from the USA in 1937. A genetic cluster, encompassing samples from the Berg (BE), Doring (DO), Kouga (KO), and Olifants (OL) Rivers, suggest shared ancestry with the H.N USA samples (Figure 2.6), but the remainder of the invasive South African populations belong to four additional clusters, hinting at the idea of multiple introductions. The ABC results corroborate this notion, as the best-fit scenario suggested a more recent introduction from the USA. However, when considering the invasive South African individuals associated with the H.N STRUCTURE cluster (Figure 2.6) as a separate South African population ($C.I_S$), the ABC analyses supported the STRUCTURE results and suggested at least two introductions: one coinciding with the recorded historic introduction and at least one more recent introduction. The observed admixture between the C.I and C.N populations (Figure 2.6) suggests that the more recent introduction also originated from the USA. Although my results support the notion of multiple introductions, this should be interpreted with caution as several factors may be responsible for this pattern, including an unsampled source population, post invasion genetic drift, insufficient marker resolution and admixture in the source population (Gray et al. 2014). Given that hatcheries make use of artificial selection techniques to enhance species production and abundance (Aprahamian et al. 2003; Lamaze et al. 2012), it is highly likely that the introduced *M. dolomieu* stocks were admixed and/ or hybrid fish, as has been

reported for stockings of salmonid fishes (Sloss et al. 2008; Cooper et al. 2010; Lamaze et al. 2012).

Invasive species capable of harbouring large, genetically diverse source populations are thought to make better invaders (Gaither et al. 2013), as they are equipped with high adaptive potential (Dlugosch 2006; Lavergne & Molofsky 2007; Wellband & Heath 2017), ultimately assisting in local adaptation. Within the invasive South African range, *M. dolomieu* experiences an array of climatic conditions with fluctuating rainfall and temperature regimes (Rutherford et al. 2006). However, despite this, *M. dolomieu* has not only survived, but has established and spread throughout the systems into which it was introduced (de Moor & Bruton 1988). Although the initial introduced individuals may have been of admixed stock, as mentioned before, the substantial amount of admixture observed within *M. dolomieu* may be due to hybridisation events within the invaded range post-introduction, as observed in *M. dolomieu* introductions elsewhere (Whitmore & Butler 1982; Whitmore & Hellier 1988; Avise et al. 1997; Pipas & Bulow 1998; Bagley et al. 2011). In addition, although sampling was conducted away from known angling 'hotspots', the human-mediated spread and 'mixing' of *M. dolomieu* cannot be ruled out.

Molecular techniques are indispensable tools in invasion biology (Muirhead et al. 2008; Blanchet 2012), particularly when wanting to reconstruct a species invasion route and/ or history (Wilson et al. 2009; Estoup & Guillemaud 2010). However, sampling problems such as the number of native versus invasive populations sampled and/ or the number of individuals sampled per population, may hinder the accuracy of the molecular markers to identify the source population (Muirhead et al. 2008). To date, however, no study to my knowledge has looked at the effect that 'sampling locality' may have on each populations' genetic composition, and hence, genetic diversity. For example, aquatic freshwater species, in particular fish, are often collected from natural or man-made reservoirs due to the ease of collection and the large number of individuals present. These specific sampling sites, however, often display much lower levels of genetic diversity when compared to rivers, as indicated by my results (localities BU and MP in invasive range; Table 2.6, Figures 2.4, 2.6). To illustrate, a recent study by Hargrove et al. (2017), reconstructing the invasion history of *M. salmoides*, identified extremely low levels of neutral genetic diversity within the invasive populations. However, much of the sampled localities in the invaded range were dams. Their results revealed that all dam populations had allele frequencies dominated by a single allele, but this

was not the case for the weir population. Thus, caution should be taken when inferring a populations' demographic and/ or invasion history as sampling bias may affect the interpretation of the results.

In conclusion, while studies comparing contemporary genetic variation among native and invasive ranges are undeniably valuable (Lozier & Cameron 2009), incorporating historic DNA is essential as it allows one to monitor temporal changes in genetic diversity often overlooked when only comparing contemporary data (Hansen 2002; Lozier & Cameron 2009). My results corroborate this idea, proving that genetic diversity can change over time. Furthermore, by incorporating historical DNA dating back to the timeframe associated with the initial introduction, I showed that *M. dolomieu* was introduced into South Africa more than once, with the genetic signature of the first introduction still being present within various river systems. Caution should, however, be taken when working with historical specimens as the degraded nature of the DNA not only hampers the successful amplification of the specimens (Sefc et al. 2003; 2007), but also renders it susceptible to genotyping discrepancies. Despite this, I recommend future studies attempting to infer the demographic history of an invasive species to incorporate historic native samples.

CHAPTER 3

INTROGRESSIVE HYBRIDISATION BETWEEN TWO INVASIVE FRESHWATER FISH, *MICROPTERUS DOLOMIEU* AND *M. SALMOIDES*

ABSTRACT

Introgressive hybridisation (IH), particularly between native and introduced species, has been observed in an array of taxa. Freshwater fish have received considerable attention in this regard, because many behavioural, biological and ecological characteristics besides external fertilisation and genomic compatibilities are often shared among species, thereby increasing the likelihood of IH. Introgressive hybridisation is thought to be detrimental to hybrid survival, but recent studies suggest that this interspecific admixture may provide sufficient genetic variation to facilitate the establishment and spread of novel populations in the invaded range. Though many studies look at IH between native and invasive species, few have examined the IH between two invasive species in a novel invaded range, and more specifically, how introgression may facilitate the invasive success of an alien species in the novel range. Using two notorious freshwater invaders (*Micropterus dolomieu* and *M. salmoides*) as model organisms, I test the hypothesis that hybridisation and introgression can occur between two invasive species in a novel invaded range. Using nine microsatellite markers and two mtDNA gene regions, I assess the possibility of IH between *M. dolomieu* and *M. salmoides*. Despite large uncorrected pairwise distances being observed between the two species (cytb: 11 – 13 %; COI: 8 – 11 %), mitochondrial introgression was still detected. Similarly, simulation and empirical analyses revealed admixed individuals, though IH seemed to be unidirectional. My findings suggest that IH may provide a species with sufficient genetic variation to adapt to the selective pressures at hand upon introduction and support the idea that IH may play a pivotal role in the successful establishment and spread of alien invasive species.

INTRODUCTION

Hybridisation (i.e. the interbreeding between individuals from different groups or species - Harrison 1990) and introgression (i.e. incorporating genetic information from one individual/ species into another through hybridisation; Anderson & Hubricht 1938), have been added to the list of hypotheses that could explain the enhanced performance of some invasive species in a new environment (Vilà et al. 2000; Lindholm et al. 2005; Darling et al. 2014; Harrison & Larson 2014). Though uncommon and/ or confined to narrow hybrid zones (Rao & Lakshmi 1999), introgressive hybridisation (IH) between native and introduced species has been observed in various taxa, including plants (Schwarz et al. 2005; Hermansen et al. 2014), mammals (Abernathy 1994), birds (Rhymer & Simberloff 1996; Sardell & Uy 2016) and fishes (Smith 1992; Avise et al. 1997; Scribner et al. 2000; Near et al. 2003; Kovach et al. 2015). Freshwater fishes, in particular, (e.g. trout: Lariagder & Scholl 1996; Echelle & Echelle 1997; Kovach et al. 2015; carp: Lamer et al. 2010; catfish: do Prado et al. 2012; cichlids: Firmat et al. 2013; sunfishes: Whitmore & Hellier 1988; Avise et al. 1997; Near et al. 2003) have received considerable attention in this regard, because closely related species often display similar behavioural, biological and ecological characteristics, in addition to exhibiting external fertilization and genomic compatibilities (Hubbs 1955; Smith 1992; Mallet 2005; Kovach et al. 2015), thereby potentially increasing the likelihood of IH.

Hybridisation and introgression between native and introduced species can lead to a multitude of outcomes, such as hybrid swarming (Allendorf & Leary 1988; Mallet 2005), competitive exclusion, niche displacement, a decrease in the native species' fitness or even population extinction (Mooney & Cleland 2001; Mallet 2005), depending on the viability and reproductive success of the hybrid offspring. In addition, new hybrids often face small population sizes and only occur at one or a few locations, making them susceptible to backcrossing (Scribner et al. 2001; Mallet 2005; Sardell & Uy 2016). Hence, the hybrid survival is reduced, with only a small percentage becoming established and even fewer becoming widespread (Thomas 2015). Yet this interspecific admixture may provide sufficient genetic variation, either through increased genetic diversity or via the transfer of adaptive alleles, to facilitate the establishment and spread of novel populations (Wilson & Bernatchez 1998; Twyford & Ennos 2012; Berthouly-Salazar et al. 2013; Firmat et al. 2013; Kovach et al. 2015; Lowe et al. 2015; Pfenning et al. 2016), while counteracting extinction by maintaining or increasing the hybrid populations' size (Drake 2006). This, in turn, may permit population

persistence, ultimately providing larger populations and more time for local adaptation to take place (Lowe et al. 2015; Pfenning et al. 2016).

A study by Wilson & Bernatchez (1998) on mtDNA introgression between lake trout (*Salvelinus namaycush*) and arctic charr (*S. alpinus*) was one of the first to highlight the potential importance of IH variation. The authors suggested that interspecific mtDNA introgression may hold selective advantages for the receiving species (Wilson & Bernatchez 1998). Recently, Sardell & Uy (2016), studying *Myzomela* honeyeaters following natural dispersal across a geographic barrier, hypothesised that IH will only occur between native and introduced species when the novel environment lacks conspecific mating opportunities for the colonising species. Similarly, Mayr & Diamond (2001) stated that hybridisation should occur if the colonising species is greatly outnumbered by the native species, as observed during the initial stages of contact in the novel environment (Wilson & Bernatchez 1998; Scribner et al. 2001; von der Heyden & Connell 2012; Pfenning et al. 2016).

Though many studies look at IH between native and invasive species, few have examined the IH between two invasive species in a novel invaded range, and more specifically, how this genomic invasion (i.e. introgression) (Mallet 2005) may facilitate the invasive success of an alien species in the novel range. South Africa is considered a global freshwater fish invasion hotspot (Leprieur et al. 2008), with non-native fishes a common constituent in all major river systems (van Rensburg et al. 2011). Two of the most notorious freshwater invaders belong to the genus *Micropterus*: the smallmouth bass, *M. dolomieu* (Lacepède 1802), and the largemouth bass, *M. salmoides* (Lacepède 1802) (Ellender & Weyl 2014). Both species have a well-documented introduction and subsequent spread history (de Moor & Bruton 1988), and have been introduced to an array of water bodies throughout South Africa (de Moor & Bruton 1988; Skelton 2001), with the initial stocking of *M. salmoides* commencing in 1930, followed by *M. dolomieu* in 1938. The two species differ in their use of freshwater habitat: *M. dolomieu* prefers clear, fast flowing waters with loose rocky substrate, such as river streams, whereas *M. salmoides* favours clear, slow-flowing or stagnant waters with a high abundance of submerged and floating vegetation, as found in impoundments (Skelton 2001). However, both species are frequently introduced into the same water bodies (i.e. rivers and impoundments) for angling purposes (de Moor & Bruton 1988; Skelton 2001; Hargrove et al. 2015). A well resolved phylogeny, (based on 1140 bp of the cytochrome b (cytb) gene fragment) has been constructed for *Micropterus*, with an uncorrected pairwise genetic distance of 13 % (cytb) being observed

between *M. dolomieu* and *M. salmoides* (Bagley et al. 2011). Evidence for natural hybridisation among *Micropterus* species is sparse (but see Barthel et al. 2010), with all known hybridisation events occurring after the human-mediated introduction of one species into another species' range (Whitmore & Butler 1982; Avise et al. 1997; Alvarez et al. 2015) or when two or more species are introduced into a man-made reservoir or impoundment as observed between *M. dolomieu* and *M. salmoides* in Texas, USA (Whitmore & Hellier 1988). These hybridisation events are, however, thought to be rare, with Whitmore & Hellier (1988) calculating hybrids to represent less than 5 % of the reservoir population. Moreover, they conclude that the hybridisation event be ascribed to at least three of the five factors listed by Hubbs (1955): (1) an artificial environment (2) large population size differences (3) the introduction of a species into a novel habitat (4) the coexistence of a small number of species in a given region and (5) impoundments covered with vegetation leading to crowded spawning grounds. This background provides a novel and suitable framework to use *M. dolomieu* and *M. salmoides* as model organisms to test the hypothesis that hybridisation and introgression can occur between two invasive species in a novel invaded range. If found to be the case it could have major implications for our understanding of invasive species and the drivers facilitating their invasive success. For example, IH may promote an 'Invasional meltdown' (Simberloff & Von Holle, 1999), whereby the disturbance initiated by one invasive species assists the establishment of successive invaders (Simberloff & Von Holle 1999; Simberloff 2006; Jackson et al. 2014).

MATERIALS AND METHODS

Sampling

Twenty-two smallmouth bass (SMB) samples were collected from the Clanwilliam Dam, Olifants River system (OL), with 17 largemouth bass (LMB) being collected from three river systems in which the two species occur sympatrically, namely the Clanwilliam Dam in the OL (n = 10), Berg River (BE) (n = 5), and the Breede River (BR) (n = 2) (Figure 3.1). Largemouth bass specimens were pooled for all analyses in order to encompass as much *M. salmoides* allelic variation as possible and to have similar sample sizes for each species. All samples were collected through angling (excluding a fry which was collected with a hand-held net), in the austral summer of 2015 – 2016 (CapeNature permit number 0056-AAA043-00004; Ethical clearance reference number SU-ACUM14-00011, University of Stellenbosch). Both

species were morphologically identified using the most defining characteristic separating the two species, namely the position of the mouth when closed (Figure 3.2 - Skelton 2001). Upon collection, a piece of muscle tissue and/or pelvic fin clipping was taken and stored in 95% ethanol for further DNA analysis.

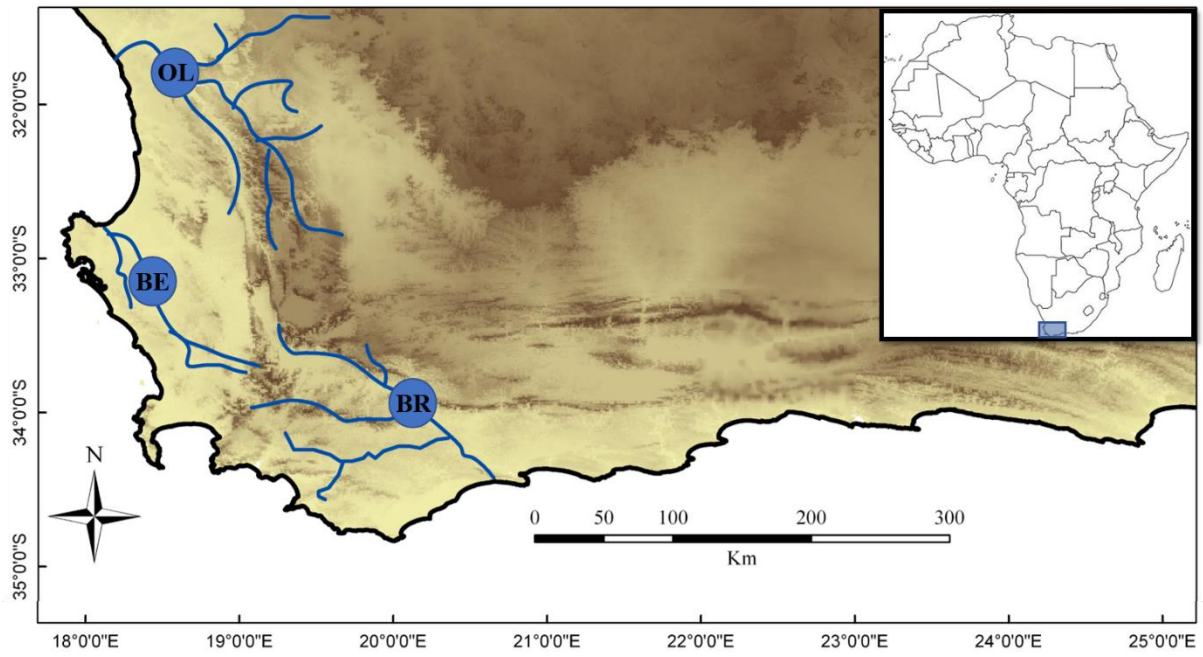


Figure 3.1. Sampling localities for *M. dolomieu* and *M. salmoides* in the present chapter. The abbreviations OL, BE and BR denotes the Olifants River, Berg River and Breede River, respectively.

DNA extraction and amplification

Total genomic DNA was extracted from each tissue sample using the NucleoSpin Tissue extraction kit (MACHEREY-NAGEL, Separations, Cape Town, South Africa) following the manufacturers protocol. To validate the morphological identification of the collected individuals, a portion of two mitochondrial (mtDNA) gene regions, namely cytochrome b (cytb) and cytochrome oxidase subunit I (COI) was amplified through a polymerase chain reaction (PCR). For the 25 µL cytb PCR reaction, the following mix was used: 01 µL Supertherm Taq polymerase, 2.5 µL buffer, 3 µL $MgCl_2$ (all three supplied by JMR Holdings), 0.5 µL dNTP mix (10 mM), 0.5 µL each of forward and reverse primers (10 mM), 14.9 µL ddH_2O and 2-5 µL genomic DNA. The internal cytb primers, basscytbf1 (5´-CAC CCC TAC TTC TCC TAC AAA GA- 3´) and basscytbr1 (5´-AAG GCR AAG AAG CGG GTG AGG G- 3´; Near et al. 2003) were used to amplify the cytb fragment under the following PCR thermocycling profile: denaturation at 95 °C for 1 min, followed by 40 cycles

of 40 s at 95 °C, 40 s at 56 °C (annealing temperature) and 1 min at 72 °C, with the final extension performed for 10 min at 72 °C. Similarly, a 20 µL PCR mix was used for amplifying the COI gene fragment, using the standard barcoding primers, VF2 (5´-TCA ACC AAC CAC AAA GAC ATT GGC AC- 3´) and FishR2 (5´-ACT TCA GGG TGA CCG AAG AAT CAG AA- 3´; Ward et al. 2005). The PCR mix contained the following: 01 µL Taq polymerase (5 U/ µL, TaKaRa TaqTM), 2 µL of 10x Buffer (Mg^{2+}), 1.6 µL dNTP mix (2.5 mM) (both from TaKaRa TaqTM), 0.4 µL of both forward and reverse primers (10 mM), 6.25 µL ddH2O, 6.75 µL sugar and 2-5 µL genomic DNA. The PCR cycling conditions were as follows: denaturation for 2 min at 94 °C, then 35 cycles of 30 s at 94 °C, 40 s at 49 °C (annealing temperature) and 1 min at 72 °C, with the final extension performed for 10 min at 72 °C. Amplified PCR products were sequenced on an automated sequencer (ABI 3730 XL DNA Analyzer, Applied Biosystems, CAF, Stellenbosch, South Africa) before being visually inspected and aligned in Geneious® 10.0.2 (Biomatters, Auckland, New Zealand).

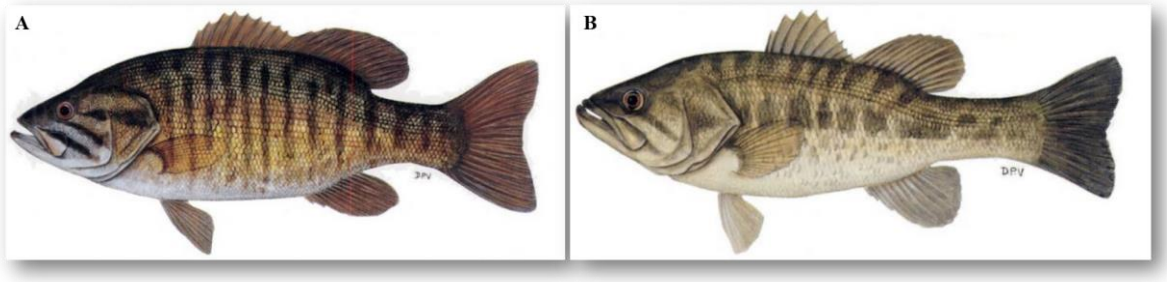


Figure 3.2. Images taken from Skelton (2001) of a smallmouth bass (*M. dolomieu*) (A) and a largemouth bass (*M. salmoides*) (B) illustrating the morphological differences between the two species.

To test for the presence of hybrid individuals, nine nuclear microsatellites markers (Mdo3, Mdo4, Mdo5, Mdo7, Mdo8, Mdo9, Mdo10, Mdo11, Lma21 - Colbourne et al. 1996; Malloy et al. 2000) were genotyped in three multiplex reactions (Table 3.1), with the following conditions for each 10 µL reaction: 5 µL Kapa2G™ Fast Multiplex Mix (Kapa Biosystems, Cape Town, South Africa), 0.2 µL of each primer (10 nmol) and 1 µL template DNA. The PCR cycling parameters followed those specified under the Low Plex cycling step, as specified by Kapa2G™ Fast Multiplex PCR Kit: denaturation for 3 min at 95 °C, followed by 30 cycles of 15 s at 95 °C, 30 s at primer dependant annealing temperature (see Table 3.1) and 30 s at 72 °C, with the final extension being performed for 10 min at 72 °C. All microsatellite genotyping was performed on the ABI 3730 XL DNA Analyzer (Applied Biosystems, CAF, Stellenbosch,

South Africa), using LIZ as an internal size marker. Geneious® 10.0.2 (Biomatters, Auckland, New Zealand) was used to score the microsatellites. To ensure accurate scoring, reference individuals previously scored were used as positive control.

Mitochondrial DNA analyses

To assess the population structure within and between the two species, a NeighbourNet Network was constructed for both the cytb and COI gene fragments using Splitstree 4.10 (Bryant & Moulton 2004). ARLEQUIN version 3.5.2.2 (Excoffier & Lischer 2010) was used on the pooled datasets (SMB and LMB) to calculate the haplotype (h) and nucleotide (π) diversity of each species and gene fragment.

Microsatellite genotyping and analyses

Genetic diversity and population structure

Microsatellite markers were tested for Hardy-Weinberg equilibrium and linkage disequilibrium in Genepop 4.2.1 (Rousset, 2008), before being assessed for amplification errors associated with large allele drop out, the presence of null alleles and stuttering in MICROCHECKER (Van Oosterhout et al. 2006). As the majority of loci for both *M. dolomieu* and *M. salmoides* were found to not be in Hardy-Weinberg equilibrium and null alleles were detected for Mdo3, Mdo7, Mdo8 and Lma21 in the *M. salmoides* dataset, FreeNA (Chapuis & Estoup 2007) was employed to re-assess the presence of null alleles using the EM algorithm (Dempster et al. 1977) for each population and locus. Null allele frequency values below 20% have been shown to not significantly impact estimation of genetic divergence (Chapuis & Estoup 2007), and was consequently used as the cut-off value.

Table 3.1. Microsatellite loci amplified in the present study with the corresponding primer sequence, size, repetition pattern, optimised annealing temperature and dye labels used.

Loci	Primer Sequence	Reference	Size (bp)	Repetition Pattern	Ta (°C)	Multiplex #	Dye Label
Mdo3	F: 5' AGGTGCTTTGCGCTACAAGT 3' R: 5' CTGCATGGCTGTTATGTTGG 3'	Malloy et al. 2000	135	$(CA)_{20}$	53.9	1	6-FAM
Mdo4	F: 5' TCTGAACAACTGCATTTAGACTG 3' R: 5' CTAATCCCAGGGCAAGACTG 3'	Malloy et al. 2000	142	$(CA)_{11}$	53.9	1	NED
Mdo5	F: 5' CAGGTTCCCTCTCACCTTCA 3' R: 5' ATGGTCTCACCAGGGACAAA 3'	Malloy et al. 2000	200	$(CT)_gCC(CA)_{10}GA(CA)_3TA(CA)_2$	61.0	3	PET
Mdo7	F: 5' TCAAACGCACCTTCACTGAC 3' R: 5' GTCACTCCCATCATGCTCCT 3'	Malloy et al. 2000	172	$(CA)_{12}$	53.9	1	VIC
Mdo8	F: 5' GTGAGGACCAGCCAAAATGT 3' R: 5' GGAAGATTGAGGTCCCAACA 3'	Malloy et al. 2000	220	$(CA)_{19}$	58.3	2	NED
Mdo9	F: 5' TTTGATGGGCGTTTTGTGTA 3' R: 5' GACCGGTCCTGCATATGATT 3'	Malloy et al. 2000	126	$(GT)_{10}$	58.3	2	PET
Mdo10	F: 5' GTGTCTCCGTGTGTTGATGG 3' R: 5' AACACCAGAGGCAAACAAGC 3'	Malloy et al. 2000	101	$(GT)_{10}$	58.3	2	VIC
Mdo11	F: 5' TTGTGGAGAGGGGCATAAAC 3' R: 5' GCATCCTCCCACGTTACCTA 3'	Malloy et al. 2000	174	$(GT)_{11}GA(GT)_3$	58.3	2	6-FAM
Lma21	F: 5' CAGCTCAATAGTTCTGTCAGG 3' R: 5' ACTACTGCTGAAGATATTGTAG 3'	Colbourne et al. 1996	158-182	$(TC)_{19}(AC)_{11}$	61.0	3	6-FAM

To determine if the individuals of each species clustered into one group, STRUCTURE 2.3.4 (Pritchard et al. 2000) was used on the microsatellite data. Ten preliminary runs for K = 2 were performed, using the no-admixture model (with independent allelic frequencies). The burn-in was set to 50,000 followed by 250,000 MCMC iterations. Structure Harvester (Earl 2012) was then used to collate the STRUCTURE results and to determine the most probable K, based on the Evanno method (Evanno et al. 2005), before exporting the appropriate input files required by CLUMPP (Jakobsson & Rosenberg 2007). Distruct (Rosenberg 2004) was used to visualise the final results.

To determine whether the microsatellite dataset was adequate to distinguish between *M. dolomieu* and *M. salmoides*, pairwise F_{ST} values, using the ENA correction and assessing the statistical significance with 10,000 iterations, was calculated in FreeNA (Chapuis & Estoup 2007). Secondly, a factorial correspondence analysis (FCA) as implemented in GENETIX (Belkhir et al. 1996), was used to visualise the genetic variation across all specimens. This method employs allelic frequencies and genetic relatedness to group individuals, rather than Hardy-Weinberg equilibrium assumptions. Lastly, the distribution of allelic frequencies for each microsatellite locus was plotted to evaluate their presence and frequency within each species. As homoplasy may mimic hybridisation through insertions or deletions in the flanking region or core sequence of the microsatellite (Estoup et al. 2002; Ellegren 2004; Henriques et al. 2016), the loci displaying allelic overlap between the two species were sequenced for six individuals of each species in order to compare their flanking regions, and so determine if the observed admixture is due to hybridisation, incomplete lineage sorting or homoplasy (Henriques et al. 2016). Lastly, genetic diversity indices, such as allelic richness (AR), number of alleles (Na), observed heterozygosity (H_O), expected heterozygosity (H_E) and the inbreeding coefficient (F_{IS}) were calculated in FSTAT (Goudet 1995), Genotype Viewer (S. Kalinowski, unpublished) and Genepop (Rousset 2008) for each species.

Hybridisation events

To test the hypothesis that hybridisation occurs between *M. dolomieu* and *M. salmoides*, a two-fold approach was implemented. Firstly, to determine the suitability of the microsatellite dataset in accurately detecting multiple hybridisation events, HybridLab (Nielsen et al. 2006) was used to simulate five hybrid states: “pure” species, F1, F2, backcrosses of F1 with *M. dolomieu* and backcrosses of F1 with *M. salmoides*. A total of 40 hybrids were simulated for

each state. Preliminary mtDNA analyses identified two of the 22 *M. dolomieu* specimens as being *M. salmoides* (though both specimens clustered with the 20 *M. dolomieu* in the STRUCTURE analysis), whereas one of the *M. salmoides* specimens, based on the microsatellite dataset, clustered with the *M. dolomieu* specimens. Hence, these three individuals were excluded from the dataset used to generate the "pure" species (thus n = 20 *M. dolomieu* and 16 *M. salmoides*) genotypes. To test the accuracy of STRUCTURE (Pritchard et al. 2000) in detecting different levels of hybridisation, all five datasets (simulated genotypes as well as confirmed "pure" species genotypes) were run using the admixture model with independent allele frequencies. Five replicates of K = 2, with an initial burn-in of 50,000 and the same number of MCMC steps, were conducted on each dataset. Following Henriques et al. (2016) the posterior probability of assignment (*q*) was used to identify the putative hybrids. A consensus is yet to be reached regarding the value of *q* to be used to distinguish between "pure" and admixed individuals (Vähä & Primmer 2006), as *q* = 0.1 (10 % admixture) decreases the chances of incorrectly assigning "pure" individuals to hybrid status, and *q* = 0.2 (20 % admixture) increases the accuracy to correctly assign individuals to their classes (i.e. "pure" vs hybrids) (Vähä & Primmer 2006; Henriques et al. 2016). Therefore, I followed Henriques et al. (2016), and set the threshold for "pure" *M. dolomieu* to *q* = 0.9 and "pure" *M. salmoides* to *q* = 0.1 (Vähä & Primmer 2006), followed by a threshold of *q* = 0.8 for "pure" *M. dolomieu* and *q* = 0.2 for "pure" *M. salmoides*. Hybrid individuals were identified as $0.1 < q < 0.9$ and $0.2 < q < 0.8$, respectively. CLUMPP (Jakobsson & Rosenberg 2007) and Distruct (Rosenberg 2004) was used to visualise the results.

Secondly, NewHybrids (Anderson & Thompson 2002), implementing a Bayesian approach, was used to corroborate the STRUCTURE results. Unlike STRUCTURE which uses Hardy-Weinberg equilibrium to detect genetic clusters, NewHybrids relies on genotypic frequencies to assess an individuals' hybrid status, thus making it more robust to hybrid detection (Anderson & Thompson 2002). Due to this, the posterior probability (*qi*) threshold was lowered to 0.5, as suggested by Aboim et al. (2010). PGDSpider 2.1.0.3 (Lischer & Excoffier 2012), along with the genotypes simulated for each of the five states ("pure" species, F1, F2, backcrosses of F1 with *M. dolomieu* and backcrosses of F1 with *M. salmoides*) in HybridLab (Nielsen et al. 2006), was used to generate the input files (for the same five states) for NewHybrids. The Jeffreys Prior was selected for both mixing and allelic frequencies, and each analysis started with an initial burn-in of 50,000 MCMC steps, followed by 200,000 MCMC iterations. Once again, each state was re-ran five times before averaging the posterior

probability over the five runs. Following the preliminary analyses, the original dataset (n = 22 *M. dolomieu* and 17 *M. salmoides*) was assessed for hybridisation events and hybrid status in STRUCTURE and NewHybrids, using the same run parameters as before.

RESULTS

Mitochondrial DNA analyses

A total of 22 *M. dolomieu* individuals, representing the OL drainage system, and 17 *M. salmoides* individuals, representing three drainage systems (OL, BE, BR) where the two species co-occur, were sequenced for a 237 bp fragment of cytb and a 409 bp fragment of the COI gene. Two of the *M. dolomieu* specimens (SMB7, SMB13) possessed mtDNA gene regions associated with that of *M. salmoides* (Figure 3.3), and were thus excluded when calculating the diversity indices for *M. dolomieu*. Eight haplotypes were retrieved for *M. dolomieu* with the cytb gene fragment, while COI rendered 16 haplotypes. Similarly, 12 and 13 haplotypes were retrieved for *M. salmoides* using the partial cytb and COI gene fragments, respectively (Table 3.2). Both partial gene fragments confirmed the distant relationship between these two species, with the uncorrected sequence divergence ranging from 11 - 13 % for cytb, and 8 - 9 % for COI. This relationship is also reflected in the NeighbourNet network (Figure 3.3). Both *M. dolomieu* and *M. salmoides* showed very similar haplotype diversity levels, but *M. dolomieu* had a much lower nucleotide diversity (Table 3.2).

Table 3.2. Genetic diversity levels for *M. dolomieu* and *M. salmoides* COI and cytb partial gene regions: number of individuals (**n**), number of haplotypes (**H**), haplotype diversity (***h***), nucleotide diversity (***π***).

Species	Gene Region	n	H	*h*	*π*
M. dolomieu	COI	20	16	0.968 ± 0.028	0.066 ± 0.034
	Cytb	20	8	0.546 ± 0.128	0.037 ± 0.020
M. salmoides	COI	16	13	0.950 ± 0.049	0.226 ± 0.115
	Cytb	14	12	0.934 ± 0.046	0.338 ± 0.172

Microsatellite genotyping and analyses

Genetic diversity and population structure

A total of 39 individuals (22 *M. dolomieu* and 17 *M. salmoides*) were successfully genotyped for nine microsatellite markers. The resulting microsatellite dataset did not conform to Hardy-Weinberg equilibrium expectations (i.e. random-mating across loci and samples), but did conform to linkage equilibrium between loci. No evidence for large allele dropout or stuttering was observed, nor were null alleles detected after the ENA correction method was employed. The STRUCTURE analysis illustrated that approximately 98% ($q \geq 0.9$) of all individuals were correctly assigned to each cluster, corroborating the species' boundaries observed with the mtDNA. Pairwise F_{ST} values, using the ENA correction, revealed no significant difference between *M. dolomieu* and *M. salmoides* in either of the loci, or across all loci ($F_{ST} = 0.452$, $P = 0.325$). The FCA analysis illustrated two distinct clusters, one for each species, with the first two axes explaining 35.30 % of all the variation. Low variation was present within *M. dolomieu* (i.e. they clustered near to one another), whereas several outlier individuals were observed for *M. salmoides* (Figure 3.4). The expected heterozygosity (H_E) per locus ranged from 0.17 – 0.75 in *M. dolomieu* and 0.11 – 0.77 in *M. salmoides*, while the observed heterozygosity (H_O) per locus ranged from 0.14 – 0.77 in *M. dolomieu* and 0.00 – 1.00 in *M. salmoides* (Table 3.3).

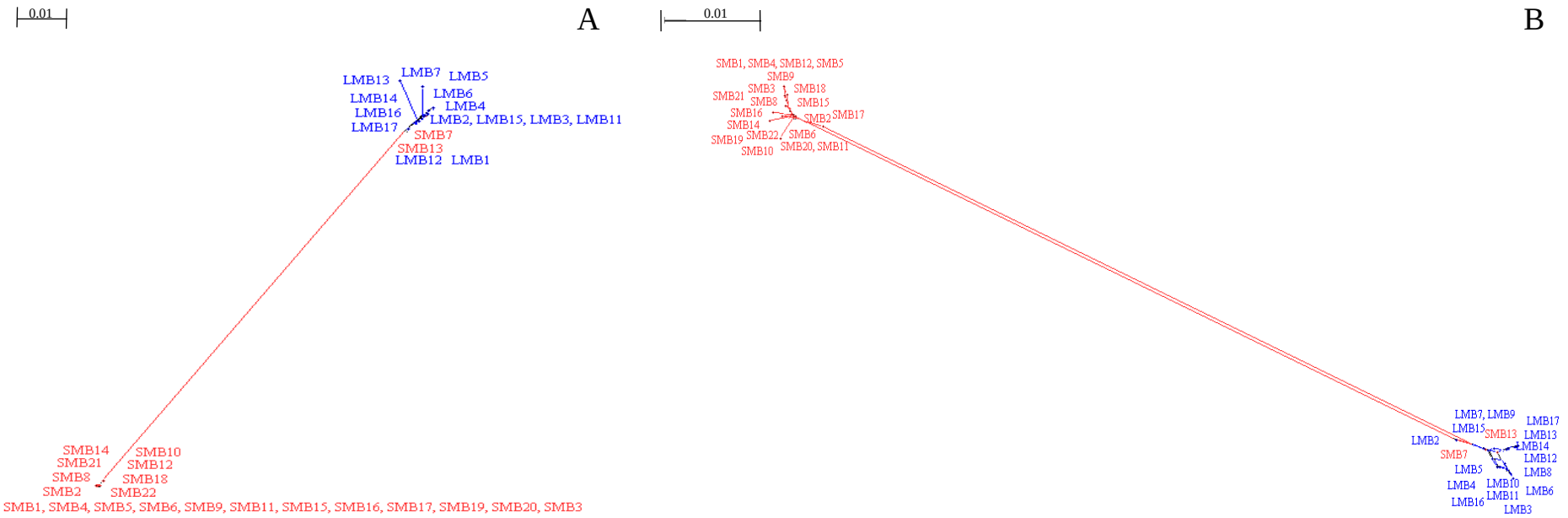


Figure 3.3. NeighbourNet network of *M. dolomieu* (red) and *M. salmoides* (blue) for the mitochondrial cytochrome b (*M. dolomieu*: *n* = 22; *M. salmoides*: n = 14) (A) and cytochrome oxidase subunit I (*M. dolomieu*: n = 22; *M. salmoides*: n = 16) (B) partial gene fragments.

Several alleles specific to either of the two *Micropterus* species could be identified from the allelic frequency histograms for the nine microsatellite loci (Figure 3.5). In particular, alleles 137 and 139 in Mdo4, 161 and 177 in Mdo7, 217 in Mdo8, 155 in Mdo9, 94 in Mdo10, 246 and 248 in Mdo11 and 173 in Lma21 appeared to be common, perhaps even fixed, in *M. salmoides*. Alleles unique to *M. dolomieu* were, however, not as common, with allele 192 in Mdo5, 169 and 171 in Mdo7 and 195, 197, 201 in Lma21 being the most abundant (Figure 3.5). No evidence for homoplasy was observed in any of the microsatellite loci when comparing *M. dolomieu* and *M. salmoides'* sequence alignment.

Table 3.3. Genetic diversity measures (**n** = number of genotyped individuals; **Na** = number of allelels; **AR** = allelic richness for a minimum of 13/17 individuals; $\mathbf{H_E}$ = expected heterozygosity, $\mathbf{H_O}$ = observed heterozygosity; $\mathbf{F_{IS}}$ = inbreeding coefficient) for nine microsatellite loci amplified for both *M. dolomieu* and *M. salmoides*.

Locus		*M. dolomieu*	*M. salmoides*
Mdo3	**n**	22	17
	Na	3.00	2.00
	AR	2.59	1.95
	$\mathbf{H_E}$	0.44	0.11
	$\mathbf{H_O}$	0.50	0.00
	$\mathbf{F_{IS}}$	-0.14	1.00
Mdo4	**n**	22	17
	Na	3.00	4.00
	AR	2.53	3.52
	$\mathbf{H_E}$	0.17	0.27
	$\mathbf{H_O}$	0.18	0.24
	$\mathbf{F_{IS}}$	-0.06	0.14
Mdo5	**n**	22	17
	Na	3.00	2.00
	AR	3.00	2.00
	$\mathbf{H_E}$	0.65	0.26
	$\mathbf{H_O}$	0.55	0.06
	$\mathbf{F_{IS}}$	0.16	0.78
Mdo7	**n**	22	17
	Na	4.00	5.00
	AR	4.00	4.72
	$\mathbf{H_E}$	0.75	0.70
	$\mathbf{H_O}$	0.77	0.41
	$\mathbf{F_{IS}}$	-0.03	0.42
Mdo8	**n**	22	13
	Na	4.00	6.00
	AR	3.57	6.00
	$\mathbf{H_E}$	0.42	0.79
	$\mathbf{H_O}$	0.36	0.23
	$\mathbf{F_{IS}}$	0.14	0.50

Table 3.3 continued on next page

Table 3.3 continued

Mdo9	**n**	22	17
	Na	4.00	4.00
	AR	2.77	4.71
	H_E	0.13	0.32
	H_O	0.14	0.18
	F_{IS}	-0.02	0.11
Mdo10	**n**	22	17
	Na	3.00	3.00
	AR	2.59	3.90
	H_E	0.53	0.22
	H_O	0.59	0.12
	F_{IS}	-0.11	-0.30
Mdo11	**n**	22	15
	Na	4.00	6.00
	AR	3.43	5.96
	H_E	0.46	0.77
	H_O	0.50	1.00
	F_{IS}	-0.09	-0.31
Lma21	**n**	22	17
	Na	4.00	5.00
	AR	3.58	4.70
	H_E	0.48	0.45
	H_O	0.55	0.24
	F_{IS}	-0.15	0.49

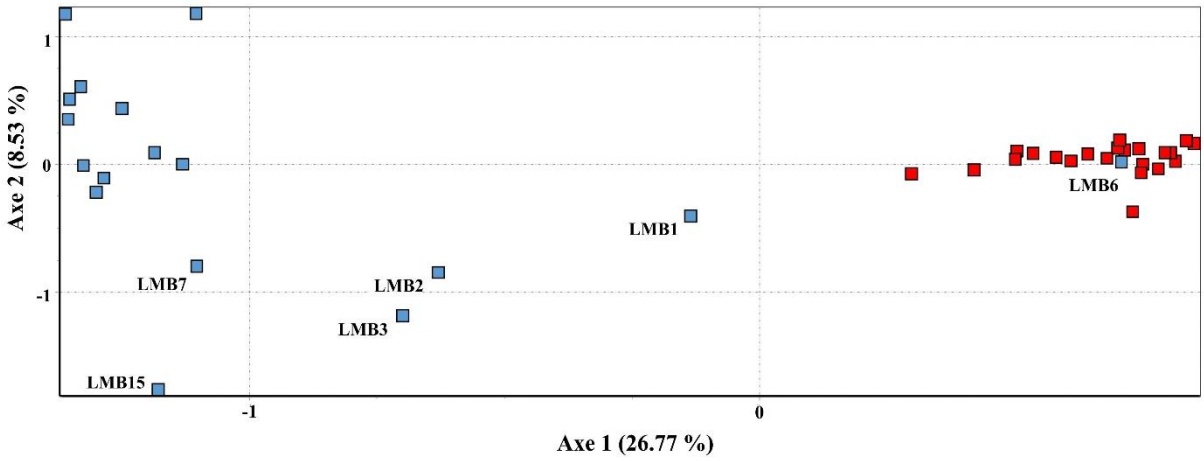


Figure 3.4. Factorial Components Analysis (FCA) for *M. dolomieu* (red) and *M. salmoides* (blue) microsatellite genotypes. Outlier individuals are labelled.

Test for hybridisation

The STRUCTURE analysis, based on the simulated genotype dataset obtained from HybridLab, revealed that accurate detection of F1 and F2 hybrid states were possible, but lost precision with an increase in hybrid states (i.e. backcrosses with either species). Detection of backcrosses was, however, drastically improved (17.5% - 20% improvement) by lowering the posterior probability of assignment (*q*) to 0.2 rather than 0.1, thus increasing the accuracy to correctly assign individuals to "pure" vs hybrid states (Figure 3.6). The NewHybrids analysis displayed an overall improvement in hybrid detection when compared to the *q* = 0.9 STRUCTURE analysis, with an overall accurate detection value of 88.75% compared to 82. 50%. Comparable results were, however obtained when compared to the *q* = 0.8 STRUCTURE analysis (overall accurate detection value of 91.88%), with the lack of accurate detection of F1 hybrids by NewHybrids decreasing its overall accuracy percentage (Figure 3.6). The original microsatellite dataset revealed two individuals, namely SMB16 and LMB1, to have admixed origins, while one individual, namely LMB6, was identified as *M. dolomieu* though both mtDNA and morphology supported its *M. salmoides* status. The NewHybrids analysis corroborated all the STRUCTURE results, in addition to identifying LMB1 as a F2 hybrid (Table 3.4; Figure 3.7). Thus, based on multiple analyses, of the 22 *M. dolomieu* individuals collected, two were identified as putative hybrids (mtDNA introgression), while two out of 17 *M. salmoides* were identified as putative hybrids (Table 3.4).

Table 3.4. *Micropterus dolomieu* (SMB) and *M. salmoides* (LMB) individuals showing evidence of introgressive hybridisation as identified through morphology, combined COI and cytb mtDNA sequence data and microsatellite genotypic assignments using STRUCTURE (along with HybridLab) and NewHybrids.

Individual	**Sampling Locality**	**Morphology**	**mtDNA**	**STRUCTURE (*q*)**	**NewHybrids**
SMB7	OL	*M. dolomieu*	*M. salmoides*	*M. dolomieu*	*M. dolomieu*
SMB13	OL	*M. dolomieu*	*M. salmoides*	*M. dolomieu*	*M. dolomieu*
SMB16	OL	*M. dolomieu*	*M. dolomieu*	Admixed (0.90): *M. dolomieu*	*M. dolomieu*
LMB1	OL	*M. salmoides*	*M. salmoides*	Admixed (0.46): *M. salmoides*	F2
LMB2	OL	*M. salmoides*	*M. salmoides*	*M. salmoides*	*M. salmoides*
LMB3	OL	*M. salmoides*	*M. salmoides*	*M. salmoides*	*M. salmoides*
LMB6	OL	*M. salmoides*	*M. salmoides*	*M. dolomieu*	*M. dolomieu*

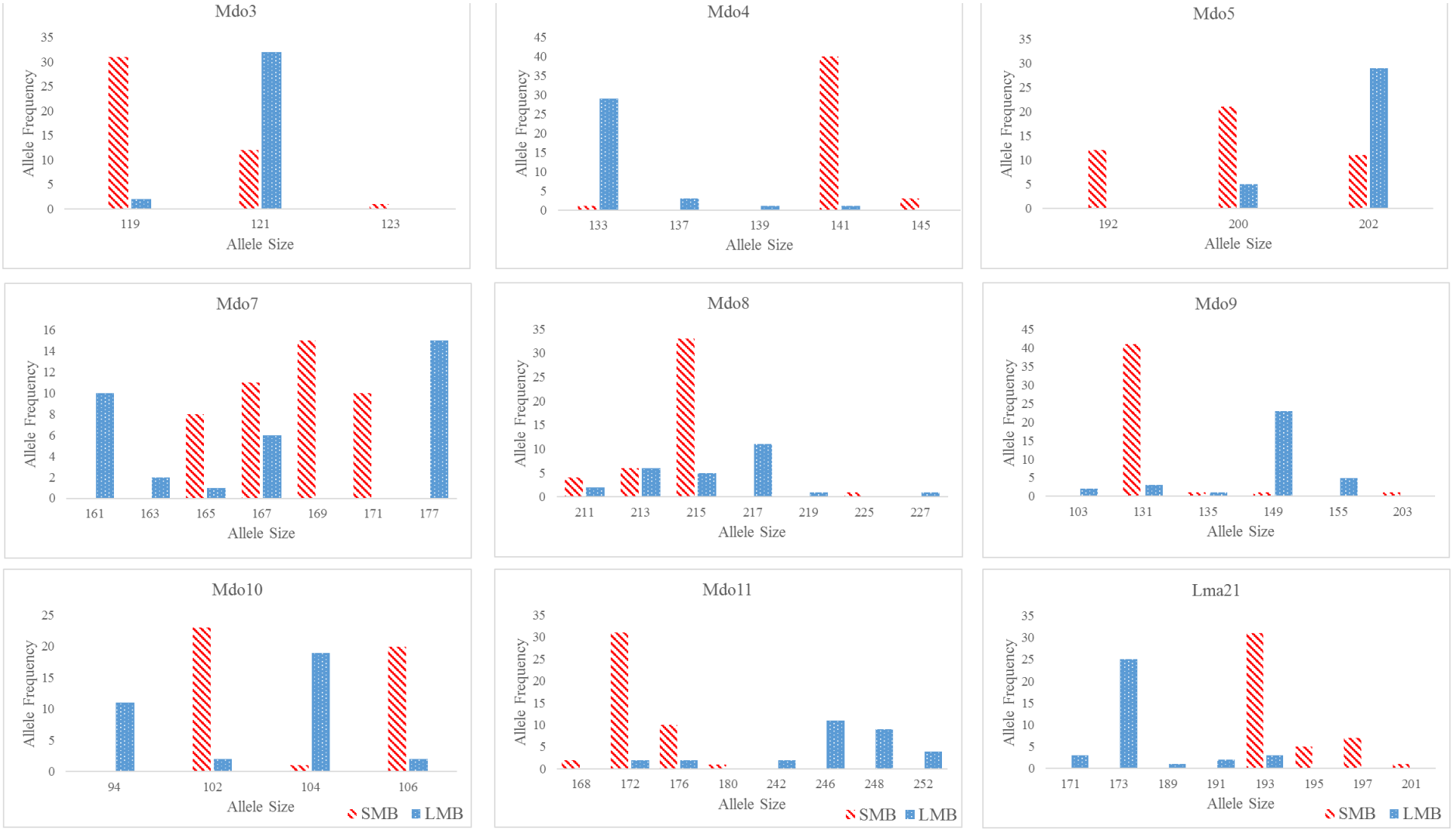


Figure 3.5. Histograms displaying the alleles and their corresponding frequencies for each of the nine microsatellite loci used in the present study. *Micropterus dolomieu* is depicted by the red bars while *M. salmoides* is represented by the blue bars.

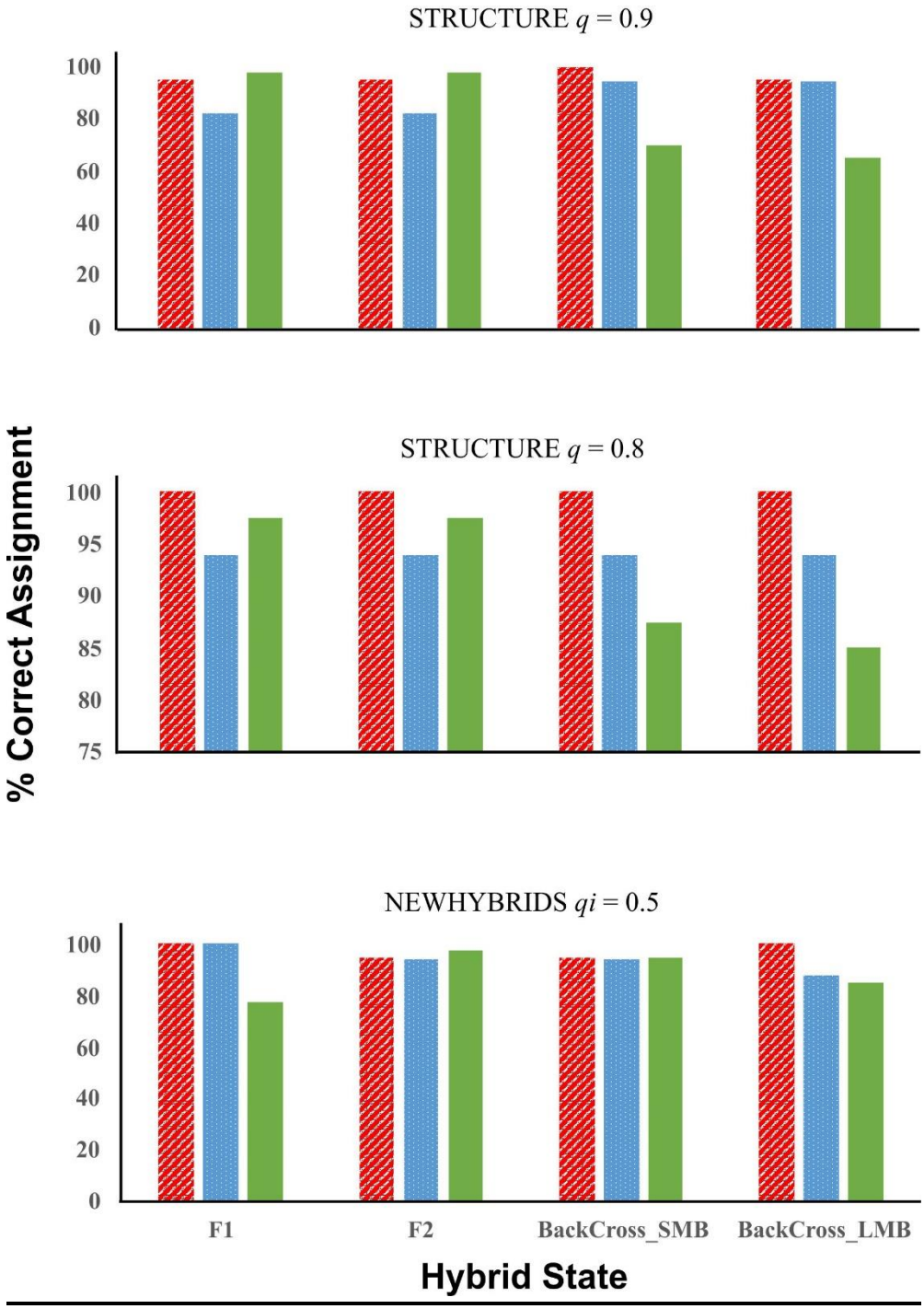


Figure 3.6. The percentage of correctly assigned hybrids to each hybrid class, as determined by STRUCTURE (based on HybridLab genotype simulations) and NewHybrids. The posterior probability of assignment (*q*) used to assess the percentage is indicated above each graph. Red bars represent *M. dolomieu*, blue bars represent *M. salmoides* and green bars denote hybrids.

.

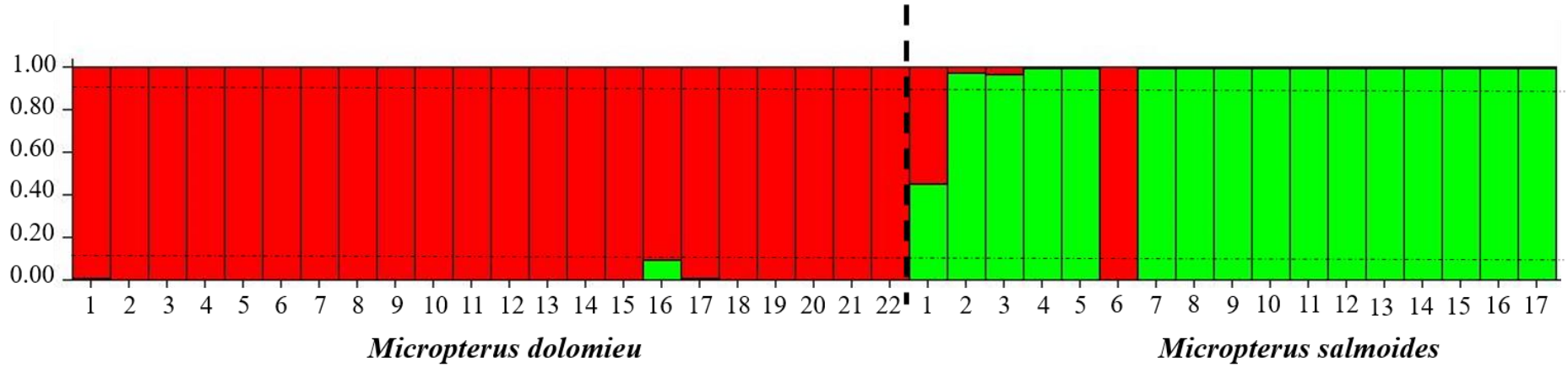


Figure 3.7. The STRUCTURE results for the original microsatellite dataset, based on multilocus assignment tests for hybrid identification (K = 2). Each red bar represents a *M. dolomieu* specimen (n = 22), while each green bar represents a *M. salmoides* specimen (n = 17). Probability of assignment to each species, as indicated by the dashed lines, was calculated by $q > 0.9$ – pure *M. dolomieu*; $q < 0.1$ – pure *M. salmoides*; $0.1 < q < 0.9$ – putative hybrids.

DISCUSSION

Introgressive hybridisation (IH), though initially thought to be rare, is increasingly being recognised as an essential source of novel genetic variation, particularly among fishes (Hubbs & Bailey 1940; Hubbs 1955; Smith 1992; Wilson & Bernatchez 1998; Kovach et al. 2015). The result of my study corroborates this trend and shows that IH can occur between two invasive species in a novel invaded range. *Micropterus dolomieu* and *M. salmoides* are known to have overlapping native distribution ranges, and are often found co-occurring both in native and introduced/invasive ranges through niche (habitat/food/resources) partitioning (Olson & Young 2003; Olson et al. 2003). Though hybridisation among *Micropterus* species is considered rare under natural conditions, introgression within this genus has been observed between an array of species, most notably when either species is introduced into an area in which it does not naturally occur (Near et al. 2003), or the introduced species' numbers are relatively small when compared to the receiving population (Hubbs 1955; Whitmore & Hellier 1988). Surprisingly, however, is that *M. dolomieu* is involved in the majority of *Micropterus* hybridisation events (Whitmore & Hellier 1988). Recorded hybridisation events within the *Micropterus* genus include: sister species, *M. dolomieu* and *M. punctulatus* (spotted bass) (Avise et al. 1997; Bagley et al. 2011), *M. salmoides* X *M. floridanus* (Florida bass) (Bagley et al. 2011) as well as *M. punctulatus* X *M. cataractae* (shoal bass) (Alvarez et al. 2015), *M. dolomieu* X *M. treculii* (Guadalupe bass) (Whitmore & Butler 1982) and *M. dolomieu* X *M. salmoides* (Whitmore & Hellier 1988). Potential IH was also documented for *M. dolomieu* X *M. coosae* (Pipas & Bulow 1998), but due to only one allozyme marker being used in this study, this IH event seems highly questionable, especially given the fact that no genus level studies since (Near et al. 2003; Bagley et al. 2011) have detected this potential crossing.

Genetic distances, (calculated for mtDNA coding regions) based on uncorrected pairwise differences (p), are often used to determine the success and viability of hybrids (Edmands 2002; Mallet 2005), with a genetic distance between 5 – 10 % being considered the threshold for freshwater systems. After this, hybrids are thought to no longer be viable (Mallet 2005). My results revealed an uncorrected pairwise genetic distance of 11 - 13 % for cytb (and 8 - 11 % for COI) between *M. dolomieu* and *M. salmoides*, corroborating the values obtained by Bagley et al. (2011) (Table 3.5). Although these values are higher than the proposed threshold values, and would thus be expected to obstruct viable offspring formation, this does not seem to be the case. Because mtDNA introgression was observed, but no F1 hybrids were

detected (Table 3.4), it's probable that this resulted from subsequent backcrossing between the hybrids and either of the "pure" *M. dolomieu* and *M. salmoides* species. This hypothesis is further supported by the fact that STRUCTURE loses detection accuracy to detect hybridisation events as the number of backcrossings increase (i.e. 97.5 % accuracy in detecting F1 hybrids, with a steady decrease to 70 – 65 % in backcrosses), resulting in these undetected backcrosses being genetically similar to one of the parental species (Figure 3.5). However, this may just be a statistical artefact from the markers used. Furthermore, it must be noted that the small sample size may have affected the statistical power of the analyses, but this does not influence the presence of hybrid individuals. Moreover, my results suggest a unidirectional IH between *M. dolomieu* and *M. salmoides*, with *M. salmoides* contributing the eggs and *M. dolomieu* the sperm. Strong directional bias can be attributed to differing egg size between *M. dolomieu* and *M. salmoides*, as observed in *S. alpinus* and *S. namaycush* (Wilson & Herbert 1993). Wilson & Herbert (1993) postulated that a cross between the smaller female species (*S. alpinus*) and larger male species (*S. namaycush*) would result in hybrids having caudal deformities due to a smaller egg volume. Differing egg volumes, particularly when the smaller eggs are produced by the female species, have long been known to impair hybrid success (Day 1884), and was corroborated by Philipp et al. (1983) who found that when crossing *M. salmoides* eggs with *M. dolomieu* sperm, the hybrid egg viability had the same success rate as *M. salmoides* eggs being fertilised by *M. salmoides* sperm. Contrastingly, however, *M. dolomieu* eggs, when fertilised with *M. salmoides* sperm, had much lower hatching rates and an increase in hybrid deformities (Philipp et al. 1983).

Table 3.5. Pairwise sequence divergence among *Micropterus* species based on the cytb gene region, as calculated by Bagley et al. (2011).

Species	*M. dolomieu*	*M. punctulatus*	*M. salmoides*	*M. floridanus*	*M. coosae*	*M. treculii*
M. dolomieu	0.000					
M. punctulatus	0.012	0.000				
M. salmoides	0.130	0.128	0.000			
M. floridanus	0.131	0.126	0.040	0.000		
M. coosae	0.133	0.132	0.113	0.107	0.000	
M. treculii	0.120	0.117	0.080	0.076	0.102	0.000

* Table adapted from Bagley et al. (2011).

In situ IH is more likely to occur in disturbed environments (e.g. introduced species in novel environments), especially if the introduced species' populations are small (Hubbs 1955; Wilson & Bernatchez 1998), as admixed offspring, though not ideal, may still contribute to the

species' persistence in the novel range (Lowe et al. 2015). Thus, hybridisation will assist in maintaining or increasing the population size, hence counteracting population extinction (Wilson & Bernatchez 1998; Mallet 2005; Pfennig et al. 2016). If this is indeed the case, it would allow the novel population more time and a larger population for intraspecific admixture (i.e. genetic rescue) and/ or for new mutations to arise, thus promoting local adaptation (see Chapter 5; Pfennig et al. 2016). This may have been the case with the smallmouth bass introduction to the Olifants River system (OL), and more specifically, the Clanwilliam Dam. The Clanwilliam Dam was initially stocked with *M. salmoides* in 1936 and 1937 and was considered established and abundant by 1945 – the year the first *M. dolomieu* fry were released into the upper reaches of the OL (Harrison 1963). As proposed by Wilson & Bernatchez (1998), *M. salmoides* mtDNA fixation following hybridisation could also have occurred via drift or founder effects. If both *M. dolomieu* and *M. salmoides* had small founding population sizes, or if small populations persisted in the Clanwilliam Dam for several generations, fixation of the mtDNA could have occurred quickly. Indirect support for the hypothesis of small founder population size and/or constrained population size can be found in the low introduction numbers (de Moor & Bruton 1988), the low Allelic richness (AR) values of the microsatellites, as well as the low mtDNA diversity, portraying a rapid population expansion (Table 3.2 and 2.3). Alternatively, *M. dolomieu* and *M. salmoides* could still currently be hybridising, but the lack of F1 detected generations does not support this notion. However, this may merely be the result of a limited sample size. Instead, these results suggest that the admixture obtained through IH may have provided *M. dolomieu* with sufficient genetic variation to adapt to the selective pressures at hand, either through the transfer of adaptive/ specific alleles from *M. salmoides* or through an increase in genetic diversity (Wilson & Bernatchez 1998; Mallet 2005; Rieseberg et al. 2007), and in doing so facilitated its range expansion.

In conclusion, my results not only show that IH is possible between *M. dolomieu* and *M. salmoides*, but also that IH between the two species appears to be common within the Olifants River system, thereby supporting the finding of Whitmore & Hellier (1988). Furthermore, my findings support the idea that IH may be correlated with the successful establishment and spread of alien invasive species (Lowe et al. 2015; Pfennig et al. 2016). Future studies should investigate the prevalence of IH between *M. dolomieu* and *M. salmoides* in other co-occurring systems and, more generally, gauge the extent of IH between introduced alien species to facilitate our understanding on the invasive potential these hybrids may possess.

CHAPTER 4

CAN ENVIRONMENTAL VARIATION INFLUENCE PHENOTYPIC CHANGES? ASSESSING THE INTRASPECIFIC MORPHOLOGICAL VARIATION IN *MICROPTERUS DOLOMIEU*

ABSTRACT

Phenotypic plasticity, a mechanism thought to shield small introductory populations from strong selection by rapidly remodelling and adapting the phenotype to the new optimum, is thought to be pivotal to the successful establishment and spread of invasive species in the novel environment. Due to their restricted nature, invasive riverine fish species are ideal study organisms to examine contemporary adaptive evolution and how this may contribute to biological invasions. Here, I use the invasive smallmouth bass (*Micropterus dolomieu*) as model organism to test the hypothesis that environmental variation drives morphological changes in phenotype. I predict that (1) similar environments will have morphologically convergent *M. dolomieu* phenotypes and that (2) most of the phenotypic variation observed will be driven by the rivers' flow regime. To test these predictions, specimens were collected from ten localities within the Olifants River system, Western Cape, South Africa, encompassing the Jan Dissels and Ratel tributaries and the mainstem. Analysis of linear and geometric morphometrics, as well as environmental variables, showed that fish inhabiting high-flow environments are more streamlined than those inhabiting slow-flowing environments.. Similarly, gape size was shown to vary among localities, with fish displaying a relatively smaller gape size being present in high-flow environments. This observed trend appears to be related to the prey composition associated with each environment. Thus, my results support the notion that similar environments will have convergent phenotypes and highlight the importance phenotypic plasticity may play in facilitating the successful colonisation, establishment and spread of invasive species.

INTRODUCTION

Organisms often face an array of selective agents and, consequently, the observed phenotype represents a fusion of morphological traits best suited to the environment while facing constrains such as gene flow and genetic drift (Slatkin 1987; Robinson & Wilson 1994; Langerhans et al. 2007). This is particularly prevalent in fishes, as variation in body form can be attributed to various functions such as feeding, swimming and predator avoidance (Wainwright et al. 2005; Langerhans 2009; Reid & Peichel 2010). For example, fish in high-flow environments have a more streamlined (fusiform) body shape and a higher 'steady swimming' performance than those from low-flow environments (Langerhans 2008). Steady swimming or cruising is employed to maintain a constant locomotion speed while swimming in a straight line, and to conduct routine tasks, such as patrolling, searching for food, migration and seeking favourable abiotic environments (Langerhans 2008). In contrast, fish in low-flow environments have deeper posterior bodies and caudal fin aspect ratios (capable of producing more thrust during rapid manoeuvres), lower steady swimming performance, but higher 'unsteady swimming' performance, essential for increased manoeuvrability and velocity or directional changes (Langerhans 2008). These results were corroborated by Langerhans (2009) who examined the effect of predation on steady- and unsteady swimming performance in mosquitofish (*Gambusia affinis*). Populations inhabiting areas of low predation risk revealed an increase in steady swimming performance when compared to areas of increased predation. However, the author noted that many fish morphological traits, such as fin size and shape, body form and muscle type may influence the observed locomotor performance. Likewise, a similar conclusion was drawn by Burns et al. (2009) studying guppies (*Poecilia reticulata*), stating that the patterns observed are more likely due to an interplay of factors and environmental influences.

Phenotypic plasticity, an organism's ability to change its phenotype in response to diverse environmental conditions (Pigliucci et al. 1996) while maintaining a single genotype (Pfenning et al. 2010), can either impede directional evolution by hindering selection (West-Eberhard 2003) or alternatively, accelerate adaptation in a novel environment (Robinson & Dukas 1999; Agrawal 2001; West-Eberhard 2003; Ghalambor et al. 2007; Cerwenka et al. 2014). Phenotypic plasticity has long been thought pivotal to the successful establishment and spread of species in new environments, particularly by invasive species (Losos et al. 2000; Richards et al. 2006; Collyer et al. 2007; Westley 2011). Considering that only a few

individuals are initially introduced into the novel environment, phenotypic plasticity is thought to be a mechanism potentially shielding these small introductory populations from strong selection (Agrawal 2001; West-Eberhard 2003; Collyer et al. 2007), by rapidly remodelling and adapting the phenotype to the new optimum (Lee 2002; Ghalambor et al. 2007; Firmat et al. 2012; Cerwenka et al. 2014; Lucek et al. 2014). For example, in the threespine stickleback (*Gasterosteus aculeatus*) in its invaded range, traits associated with locomotion are regulated through demographic processes, ultimately enabling the species' range expansion and invasion success (Lucek et al. 2014). Similarly, brown trout (*Salmo trutta*) in Canada were observed to alter their body shape in response to stream flow, with similar environmental habitats hosting phenotypically similar *S. trutta* (Westley et al. 2012). Likewise, in *Gambusia affinis*, phenotypic changes in the caudal regions and traits associated with burst-swimming abilities changed rapidly in response to fluctuating water velocities in the invaded range (Langerhans 2009). Thus, invasive riverine fish species are ideal study organisms to examine contemporary adaptive evolution and how this may contribute to biological invasions, because unlike other invasive flora and fauna, freshwater fish are restricted to the aquatic system in which they find themselves (Cerwenka et al. 2014; Peoples et al. 2017).

An example of such a species is the smallmouth bass (*Micropterus dolomieu*). Native to the east-central parts of the United States of America and two Canadian provinces, *M. dolomieu* is currently recognised as an invasive species in at least 12 countries worldwide, including South Africa (Loppnow et al. 2013). Initially introduced into South Africa in 1937 for angling purposes (de Moor & Bruton 1988), *M. dolomieu* has established in several rivers and man-made water bodies across the country (van Rensburg et al. 2011), and is particularly prevalent in the Olifants River system, Western Cape (van der Walt et al. 2016). Using *M. dolomieu* as a model organism, I hypothesise that environmental variation drives morphological changes in phenotype and, in turn, could promote the invasive success of alien species. I predict that (1) similar environments will have morphologically convergent *M. dolomieu* phenotypes and that (2) most of the phenotypic variation observed will be driven by the flow regime (Langerhans et al. 2003). Understanding the drivers of phenotypic trait variation will assist us in grasping how environmental variation may influence community structure (Jackson et al. 2016), while comprehending the factors promoting the invasion success of an organism would be essential to prevent future invasions in pristine environments (García-Berthou 2007; Hui & Richardson 2017).

MATERIALS AND METHODS

The study system

The Olifants River is one of the largest river systems within the Western Cape Province of South Africa, with an approximate length of 285 km and a catchment area of 46,220 km^2. Two catchment dams, namely Bulshoek and Clanwilliam Dam, constructed in 1919 and 1935 respectively, are situated within the Olifants River (Figure 4.1). Initially introduced into the upper reaches of the Olifants River system in 1943 and 1945 for recreational purposes, *M. dolomieu* has successfully invaded and established itself throughout the Olifants River system and its tributaries, including the two catchment dams (Harrison 1963; Van der Walt et al. 2016).

Sampling regime and field procedures

Approximately 20 *M. dolomieu* specimens were collected from ten localities (total sample size, n = 203) in the austral summer (January – March) of 2015. Sampling occurred at approximately 10 km intervals, encompassing the Jan Dissels tributary, main stem of the Olifants River (including the Clanwilliam Dam), and the Ratel tributary (Figure 4.1). All specimens were collected by angling (CapeNature permit number: 0056-AAA043-00004) using standardised lure sizes, before being euthanized with clove oil (Ethical clearance reference number SU-ACUM14-00011, Stellenbosch University). Lastly, GPS coordinates were recorded for each sampling locality.

Morphometric and geometric morphometric data collection

Both linear morphometric measurements and geometric morphometrics were employed to test the hypothesis that environmental disparity drives phenotypic variation in morphology. The rationale is that geometric morphometrics have been shown to be more effective when analyzing and interpreting body shape variation, while traditional morphometrics provide the statistical support for these changes (Parsons et al. 2003).

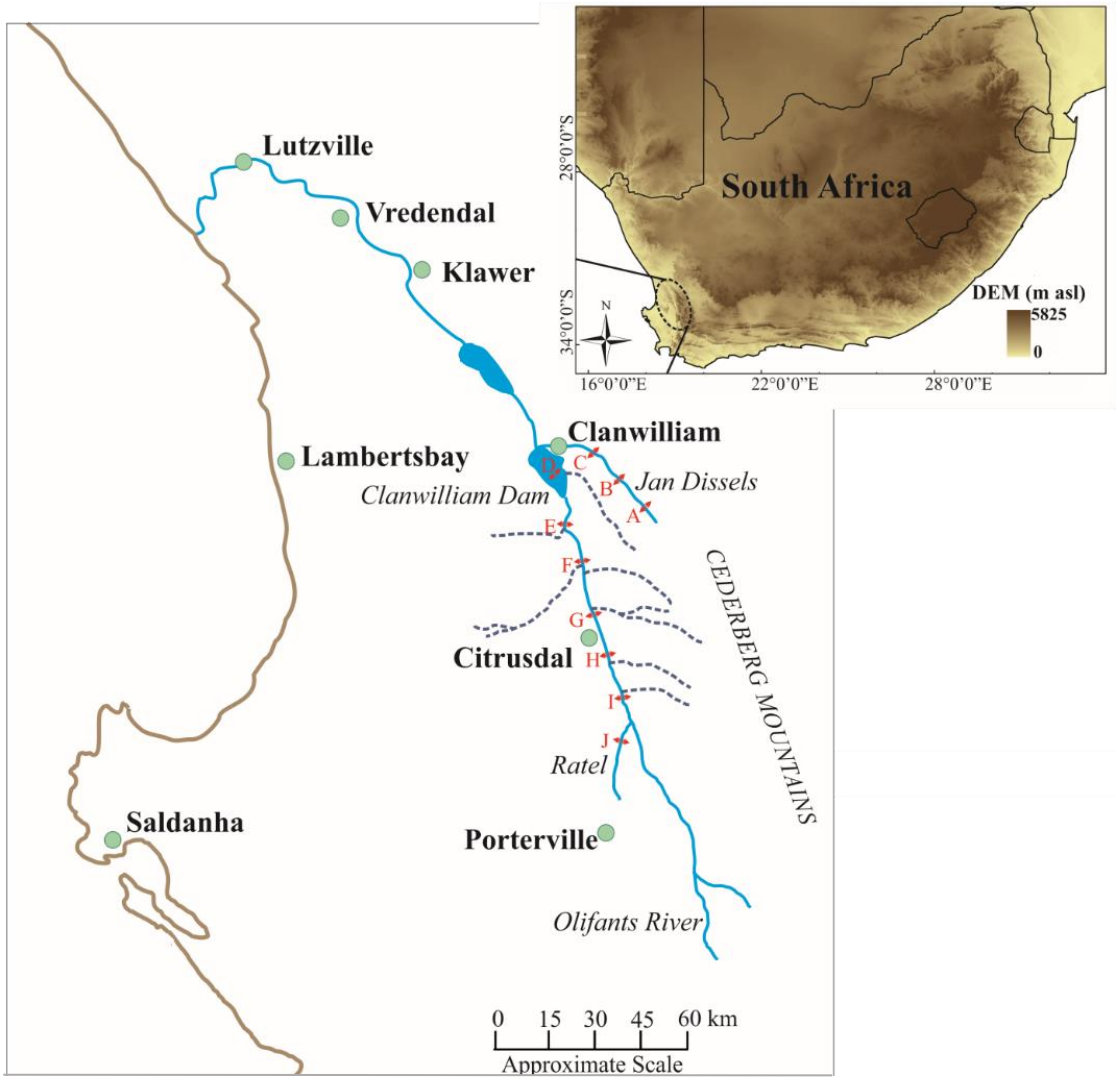


Figure 4.1. The ten sampling localities (A – J) within the Olifants River system of the Western Cape, South Africa. Localities A – C represent those sampled in the Jan Dissels tributary, while Locality D represents the Clanwilliam Dam. Localities E – I represent sampled localities situated within the mainstem of the Olifants River and Locality J represents the Ratel tributary.

Firstly, a standardized photograph of the left lateral side of each fish was taken with a digital Sony DSC-HX100V camera. To ensure standardized images, the unpreserved specimens were placed on laminated graph paper and a tripod, set to a standard height, was used. To quantify and visualise the body shape change between localities, 12 unambiguous and homologous landmarks were digitized using tpsDig2 version 2.17 software (Rohlf 2013) (Figure 4.2). Twenty-five linear morphological characters, representing both swimming and feeding traits (Winemiller 1991; Oliveira et al. 2010; Yokogawa 2013) (Figure 4.2, Table 4.1), were measured using digital callipers (precision of 0.01 mm) (Absolute Coolant Proof Caliper Series 500 IP67, Mitutoyo Ltd., Japan), before being dissected to determine the sex of each specimen. In addition, an otolith from each fish was removed, cleaned, sectioned, and the

growth rings counted to determine the age of each specimen using the procedure described by Taylor & Weyl (2013).

Environmental data

Data characterising the river topology, soil classification and hydroclimate for the Olifants River system were obtained from a near-global data set consisting of freshwater environmental variables (in a standardized 1 km grid) derived by Domisch et al. (2015). Given the potentially large spatial and temporal variation, and consequently large measurement error, localised environmental variables (e.g. pH, dissolved oxygen, salinity etc.) were omitted. For each GPS coordinate (i.e. sampling locality), 16 variables were extracted using the R packages RASTER 2.5-8 (Hijmans 2016) and SP 1.2-4 (Pebesma 2016). Firstly, four variables related to river topography were calculated. These include (1) average elevation across sub-catchment – 'ELE', (2) average river slope across sub-catchment – 'RS', (3) flow length (i.e. the number of upstream stream grid cells) – 'FL' and (4) flow accumulation (i.e. the total number of upstream grid cells) – 'FA'. Second, four variables that classify the soil type were obtained, including (5) sand content mass fraction across sub-catchment – 'SNDPPT', (6) silt content mass fraction across sub-catchment – 'SLTPPT', (7) clay content mass fraction across sub-catchment – 'CLYPPT' and (8) coarse fragments (>2 mm fraction) volumetric across sub-catchment – 'CFRVOL'. Lastly, we extracted eight bioclimatic variables, four characterising air temperature and four describing precipitation averaged across sub-catchment, which were derived from the WorldClim database (Hijmans et al. 2005). Bioclimatic variables comprise (9) annual mean temperature – 'BIO1', (10) maximum temperature of warmest month – 'BIO5', (11) temperature annual range – 'BIO7', (12) mean temperature of warmest quarter – 'BIO10', (13) annual precipitation – 'BIO12', (14) precipitation of wettest month – 'BIO13', (15) precipitation seasonality (i.e. coefficient of variation) – 'BIO15' and (16) precipitation of wettest quarter – 'BIO16' (Domisch et al. 2015). As Domisch et al. (2015) showed a strong correlation between the water- and air temperature, air temperature was used as a proxy for water temperature in the present study.

Statistical analyses - Geometric morphometrics

MorphoJ (Klingenberg 2011) was employed to analyse the geometric morphometric landmark coordinates used for studying body shape variation. Firstly, a Procrustes superimposition was performed on the 'raw' coordinates to remove orientation, size and position biases from the data (Klingenberg 2011). The resulting Procrustes coordinates were

regressed against centroid size (CS) to obtain size corrected residuals for each of the landmark coordinates. The resulting residuals were then used to construct a covariance matrix, before a principal component analysis (PCA) was performed to reduce the dimensionality of the data (hereafter referred to as 'PC_{SHAPE}'). Furthermore, a multivariate analysis of variance (MANOVA) with subsequent univariate *F*-tests was employed to compare the PC_{SHAPE} axes between localities. Subsequent Bonferroni *post-hoc* tests were conducted to determine differences in body shape between the populations.

Table 4.1. Linear morphometric measurements and their corresponding abbreviations (Abbrev). Abbreviations correspond to those used in Figure 4.2.

Trait #	Morphological Characters	Abbrev.	Trait #	Morphological Characters	Abbrev.
1	Standard body length	SBL	14	Dorsal fin height	DFH
2	Maximum body height	MBH	15	Caudal fin length	CFL
3	Maximum body width	MBW	16	Caudal fin height	CFH
4	Caudal peduncle length	CPdL	17	Anal fin length	AFL
5	Caudal peduncle height	CPdH	18	Anal fin height	AFH
6	Caudal peduncle width	CPdW	19	Pectoral fin length	PtFL
7	Head length	HL	20	Pectoral fin height	PtFH
8	Head height	HH	21	Pelvic fin length	PFL
9	Head width	HW	22	Pelvic fin height	PFH
10	Eye height	EH	23	Length of snout with mouth closed	LSC
11	Mouth height	MoH	24	Length of snout with mouth open	LSO
12	Mouth width	MoW	25	Body midpoint height	BMH
13	Dorsal fin length	DFL			

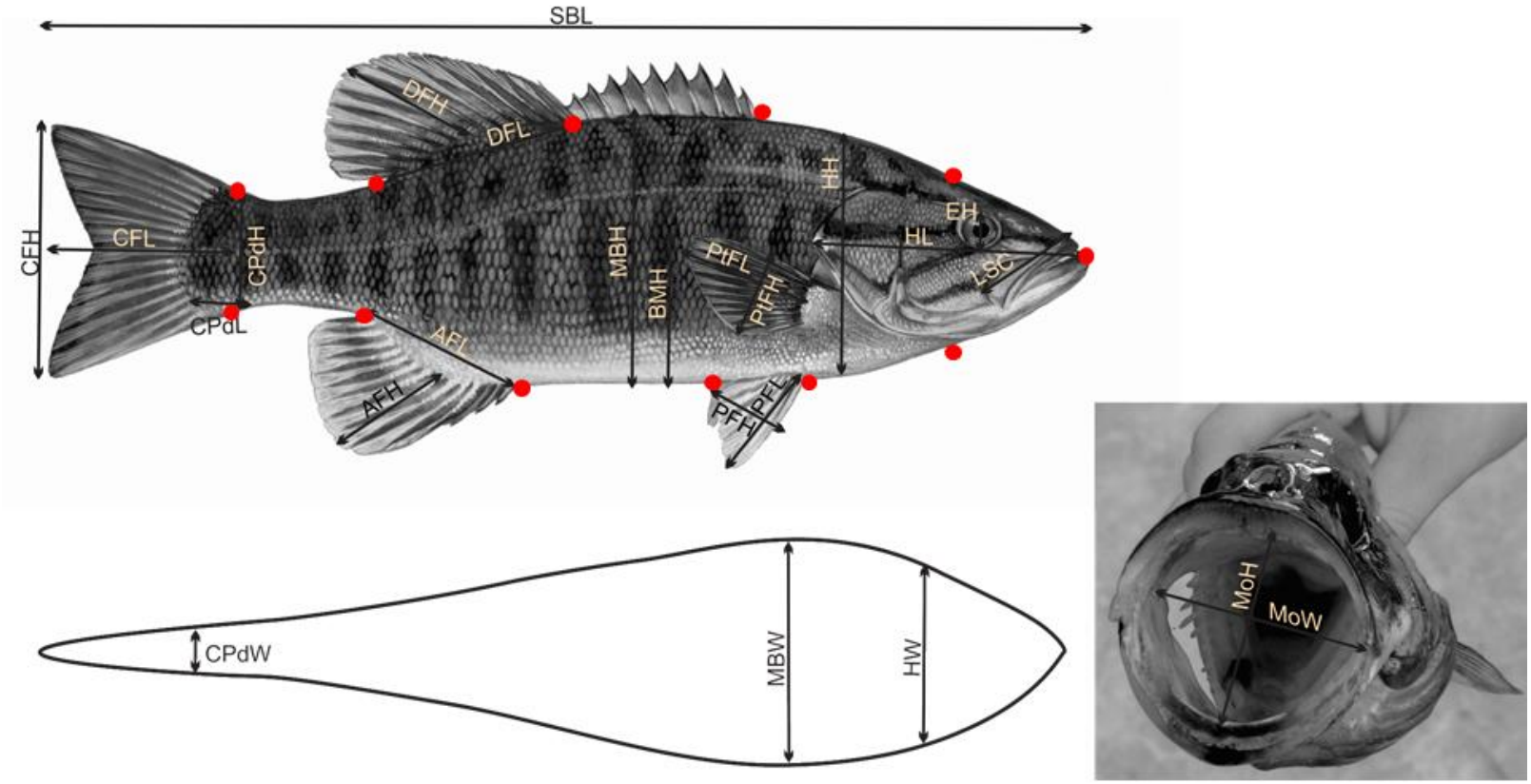


Figure 4.2. The 12 geometric morphometric measurements (●) and the 25 linear morphometric measurements taken for each specimen. Linear morphometric measurement abbreviations correspond to those in Table 4.1.

Statistical analyses - Linear morphometrics

A PCA was conducted on all the linear morphometric data, because standard body length (SBL) might not represent the actual size of the fish. Principal component 1 explained 92.8 % of the total variation (Eigenvalue = 23.2) and was retained for subsequent analyses (hereafter referred to as 'body size'). To correct for body size variation among populations, each linear morphometric variable was regressed against body size. The resulting residuals were retained for further analyses. Subsequently, a PCA was conducted on these residuals, retaining the most significant PC axes (hereafter referred to as 'PC_{LINEAR}'). The PC_{LINEAR} axes were compared between the localities using a MANOVA with subsequent univariate *F*-tests, followed by Bonferroni *post hoc* tests.

Environmental data analyses

Due to the highly correlated nature of the environmental variables, a PCA was conducted to avoid potential problems associated with multicollinearity. The PCA reduced the dataset to a smaller number of independent PC scores (hereafter referred to as '$PC_{ENVIRONMENT}$'), which were subsequently used in a linear regressionanalyses to test for the effects of environment on morphology. The PC_{LINEAR} and PC_{SHAPE} scores were used as dependent variables, whereas the $PC_{ENVIRONMENT}$ scores were used as independent variables. All statistics were done using SPSS STATISTICS v. 20.0.0 (SPSS Inc., Chicago, IL, USA).

Multi-scale pattern analysis

Lastly, to investigate and visualize the spatial structure within the data, a multi-scale pattern analysis (MSPA) (Jombart et al. 2009), as implemented in the R package *adegenet* (R Development Core Team, 2011), was conducted. This approach uses Moran's eigenvector maps (MEMs) (Dray et al. 2006) to disentangle the spatial patterns promoting ecological variability across a landscape. Named in descending order, MEM1 represents the broadest spatial scale while the last MEM describes the finest local spatial scale. The R^2 determination coefficient was used to quantify the association strength between the MEMs and the independent morphological (residual linear morphometrics and shape PC scores) and environmental variables. A K–nearest neighbor algorithm, specifying K as 20 (20 specimens were collected per population), was selected for the connection network. In addition, a redundancy analysis was employed to combine environmental variables and spatial predictors in an attempt to explain morphological variation across a landscape.

RESULTS

Size and age structure

Significant variation in size (ANOVA; $F_{9,191} = 38.37$, $P < 0.001$) and age (ANOVA; $F_{9,191} = 14.03$, $P < 0.001$) was observed among the localities. Specimens collected at Locality D were approximately twice the size of specimens belonging to the other localities (Bonferroni *post hoc* tests, all $P < 0.001$; Appendix 4.1a). Size variation among the other localities was rather limited (range 15.1 – 21.8 cm SBL; Appendix 4.1a). With regards to age, fish belonging to Locality A and D were significantly older (Bonferroni *post hoc* tests, all $P \leq 0.03$; Appendix 4.1b) than those of the other localities. No difference in age was observed among the other populations (Appendix 4.1b).

Geometric morphometrics

Four PC_{SHAPE} axes (eigenvalues $PC_{SHAPE\ 1} = 4.90$; $PC_{SHAPE\ 2} = 1.64$, $PC_{SHAPE\ 3} = 1.41$, $PC_{SHAPE\ 4} = 1.21$), explaining 67 % of the variance, were retained for the geometric morphometric data analyses. Shape PC1 ($PC_{SHAPE\ 1}$) represented fish that underwent a dorsal to ventral redistribution in body shape, while $PC_{SHAPE\ 2}$ represented fish that had rounded bodies in addition to shortened caudal peduncles (Figure 4.3). Similarly, $PC_{SHAPE\ 3}$ represented fish with wider bodies, while $PC_{SHAPE\ 4}$ denoted slender fish with large heads and shortened caudal peduncles (Appendix 4.2). The MANOVA with subsequent univariate *F*-tests revealed significant differences in body shape among the populations (Table 4.2). Moreover, Bonferroni *post hoc* tests indicated that fish from locality A had higher bodies ($PC_{SHAPE\ 1}$) than the other populations, with only marginal variation being observed with regards to locality B and D (Figure 4.3). For $PC_{SHAPE\ 2}$, $PC_{SHAPE\ 3}$ and $PC_{SHAPE\ 4}$ the significance of the univariate *F*-tests was determined by locality J, locality F and locality I, respectively.

Linear morphometrics

The principal component analysis conducted on the 25 linear morphometric measurements retained nine independent PC axes ($PC_{LINEAR\ 1\text{-}9}$) (Table 4.3). Principal component 1 ($PC_{LINEAR\ 1}$) represented fish with higher bodies (MBH), higher and narrower caudal peduncles (CPdH, CPdW), an increased head height (HH), and decreased eye height (EH), anal fin height (AFH) and pelvic fin length (PFL). Hence, $PC_{LINEAR\ 1}$ represents fish with 'deeper bodies' (Figure 4.4). Principal component 2 ($PC_{LINEAR\ 2}$) corresponded to 'gape size,' as the highest loading scores were related to mouth height and width (MoH, MoW) (Table 4.3).

The MANOVA revealed differences in all PC axes except $PC_{LINEAR\ 9}$ (Table 4.2). In addition, the Bonferroni *post hoc* tests further revealed that locality A had less deep bodies when compared to the other localities (Bonferroni *post hoc* tests, all $P < 0.05$) (Figure 4.4). No morphological differences were found between the sexes (Table 4.2). However, variation in body depth ($PC_{LINEAR\ 1}$) appeared to be related to age (Table 4.2).

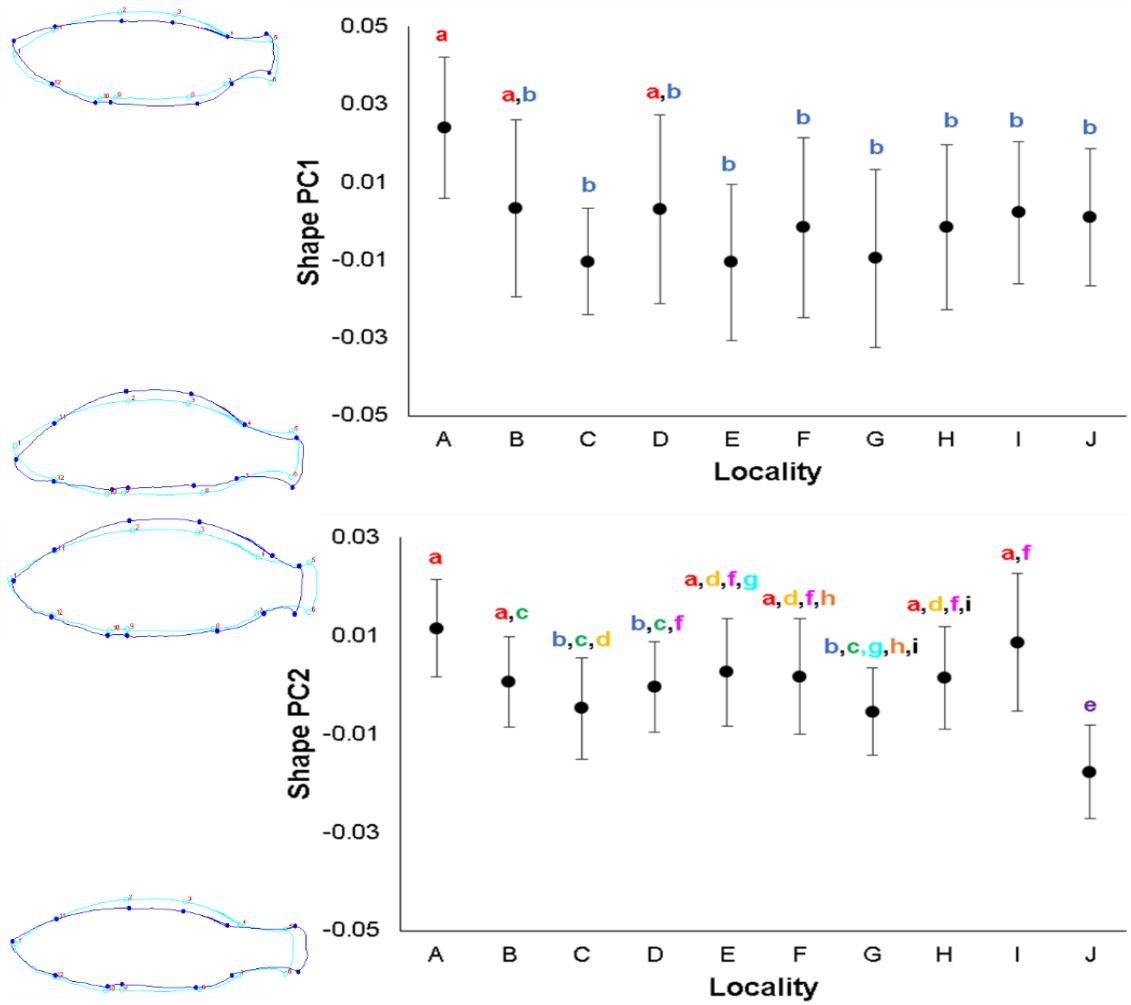


Figure 4.3. Graphs illustrating the mean and standard deviation for each locality with regards to $PC_{SHAPE\ 1}$ (dorsal-ventral redistribution) and $PC_{SHAPE\ 2}$ (rounded bodies, shortened caudal peduncle). The fish outline drawings depict the variation in fish body shape for each PC axis, with the light blue line representing the average shape for all fish, while the dark blue line represents the upper and lower body shape extremity (scale factor set to 0.08 and - 0.08, respectively). Different alphabetical letters above the bars indicate values with statistically significant differences (Bonferroni *post hoc* test; $P < 0{\cdot}05$), while identical letters indicate no significant differences among means.

Environmental data

The PCA reduced the 16 environmental traits to two independent PC axes which explained 94.8% of the total variation. The first PC axis, $PC_{ENVIRONMENT\ 1}$, was negatively correlated with elevation, slope, percentage of coarse fragments and temperature, but positively correlated with flow length, flow accumulation, precipitation and clay content (Table 4.4). Therefore, $PC_{ENVIRONMENT\ 1}$ epresents a gradient from fast-flowing streams at high altitude to wide, slow-flowing rivers at lower altitude. $PC_{ENVIRONMENT\ 2}$ was positively correlated with percentage of sand, as well as air temperature across the sub-catchment, but negatively correlated with elevation and silt content (Table 4.4). Linear regression analysis showed that environment ($PC_{ENVIRONMENT\ 1}$) had a significant effect on body depth ($PC_{LINEAR\ 1}$; standardised β = 0.77; P = 0.009) and gape size ($PC_{LINEAR\ 2}$; standardised β = -0.75; P = 0.013) (Figure 4.5) but not on the remaining size ($PC_{LINEAR\ 3\text{-}9}$; standardised β = -0.28 - 0.35; all $P \geq 0.32$) or shape measurements ($PC_{SHAPE\ 1\text{-}4}$; standardised β = -0.26 - 0.46, all $P \geq 0.10$). In contrast, $PC_{ENVIRONMENT\ 2}$ had no effect on size ($PC_{LINEAR\ 1\text{-}9}$; standardised β = -0.28 - 0.54; $P \geq 0.11$) and shape ($PC_{SHAPE\ 1\text{-}4}$; standardised β = -0.45 - 0.07; $P \geq 0.19$) measurements.

Table 4.2. Multivariate analysis of variance (MANOVA) results for both linear morphometric ($PC_{LINEAR\ 1\text{-}9}$) and geometric morphometric measurements ($PC_{SHAPE\ 1-4}$).

	Locality		**Sex**		**Age**	
	$F_{9,189}$	***P*-value**	$F_{1,186}$	***P*-value**	$F_{1,186}$	***P*-value**
$PC_{LINEAR\ 1}$	23.54	<0.001*	0.34	0.56	30.21	<0.001*
$PC_{LINEAR\ 2}$	9.21	<0.001*	1.77	0.19	6.33	0.01
$PC_{LINEAR\ 3}$	8.43	<0.001*	0.44	0.51	1.25	0.27
$PC_{LINEAR\ 4}$	4.57	<0.001*	3.42	0.07	0.25	0.62
$PC_{LINEAR\ 5}$	3.28	0.001*	2.74	0.10	0.37	0.62
$PC_{LINEAR\ 6}$	7.68	<0.001*	0.14	0.71	0.09	0.77
$PC_{LINEAR\ 7}$	2.60	0.008*	4.37	0.04	2.10	0.15
$PC_{LINEAR\ 8}$	2.87	0.003*	0.72	0.40	2.32	0.13
$PC_{LINEAR\ 9}$	1.86	0.06	3.25	0.07	0.11	0.75
	$F_{9,188}$	***P*-value**	$F_{1,188}$	***P*-value**	$F_{1,188}$	***P*-value**
$PC_{SHAPE\ 1}$	3.57	<0.001*	0.39	0.54	0.03	0.87
$PC_{SHAPE\ 2}$	10.09	<0.001*	0.29	0.59	2.34	0.13
$PC_{SHAPE\ 3}$	5.69	<0.001*	5.70	0.02	0.42	0.52
$PC_{SHAPE\ 4}$	6.79	<0.001*	0.16	0.67	1.36	0.25

* Significant after sequential Bonferroni correction

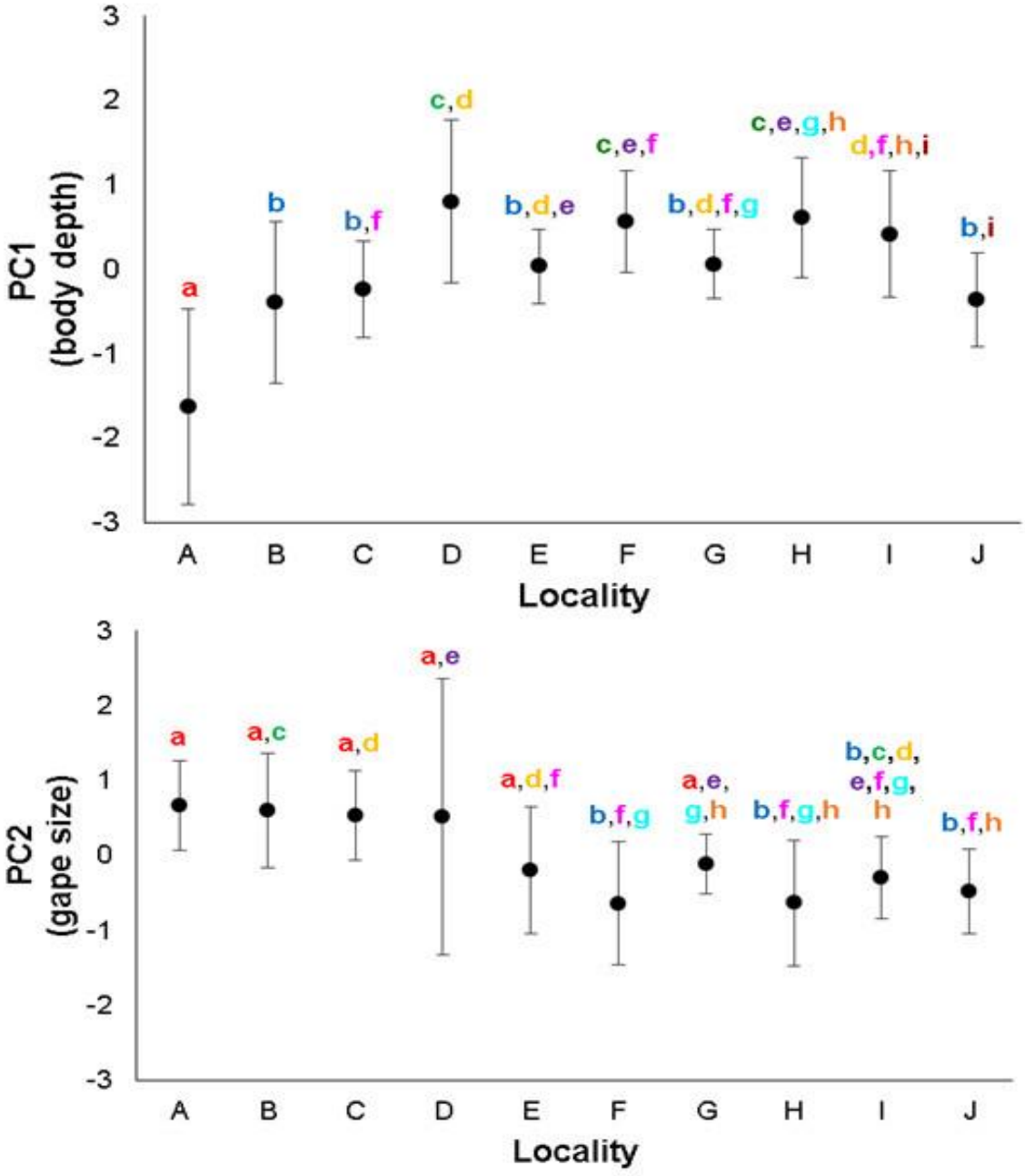


Figure 4.4. Graphs illustrating the mean and standard deviation for each locality with regards to $PC_{LINEAR\ 1}$ (body depth) and $PC_{LINEAR\ 2}$ (gape size). Different alphabetical letters above the bars indicate values with statistically significant differences (Bonferroni *post hoc* test; $P < 0{\cdot}05$), while identical letters indicate no significant differences among means.

Table 4.3. Linear morphometric PC scores ($PC_{LINEAR\ 1\text{-}9}$) for each of the 25 linear morphometric measurements. Values in bold represent loading scores greater than 0.50.

	PC1 (body depth)	**PC2 (gape size)**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**PC8**	**PC9**
SBL	0.161	0.378	0.081	0.392	0.358	-0.350	0.266	0.027	0.063
MBH	**0.539**	0.350	0.126	-0.261	0.283	-0.064	-0.042	-0.009	0.229
MBW	0.474	0.244	-0.345	-0.179	0.202	-0.340	0.134	0.044	-0.168
CPdL	0.116	0.232	-0.328	**0.562**	-0.319	0.028	-0.216	-0.317	0.032
CPdH	**0.593**	0.073	0.413	-0.041	-0.002	-0.114	0.259	-0.047	0.087
CPdW	**-0.543**	0.294	-0.175	-0.226	-0.029	-0.464	-0.205	-0.086	-0.104
HL	-0.220	-0.052	0.458	0.465	0.412	-0.021	0.156	0.105	-0.130
HH	**0.519**	0.354	0.125	0.140	0.238	0.257	-0.142	-0.146	0.138
HW	0.328	-0.178	-0.305	-0.116	0.389	-0.051	-0.089	-0.055	-0.457
EH	**-0.532**	0.058	-0.266	0.183	-0.045	0.268	0.247	-0.005	-0.142
MoH	0.424	**-0.667**	-0.084	0.192	-0.005	0.068	-0.079	0.119	0.112
MoW	0.486	**-0.637**	-0.118	0.099	0.143	0.001	-0.209	-0.053	-0.087
DFL	0.245	0.456	0.372	-0.022	-0.241	0.083	0.284	-0.016	-0.125
DFH	-0.330	-0.012	0.315	-0.346	-0.213	-0.076	0.007	0.013	-0.434
CFL	-0.211	0.139	**-0.574**	0.289	-0.175	0.074	0.317	0.120	0.215
CFH	0.415	-0.293	0.101	-0.066	-0.345	0.149	0.307	0.280	-0.262
AFL	0.176	0.313	0.123	-0.185	-0.428	-0.078	-0.209	0.452	0.179
AFH	**-0.553**	-0.211	0.259	-0.185	0.273	0.167	0.069	-0.053	0.200
PtFL	-0.486	0.141	-0.028	-0.055	0.190	0.447	0.090	0.212	0.024
PtFH	0.298	0.022	0.001	-0.406	-0.164	0.319	-0.127	-0.346	0.216
PFL	**-0.620**	-0.118	0.111	-0.200	0.221	-0.180	-0.100	-0.183	0.172
PFH	-0.044	**-0.509**	0.237	-0.111	-0.054	-0.196	0.396	-0.173	0.289
BMH	0.109	0.181	-0.229	-0.250	0.451	0.341	-0.098	0.425	0.051
LSC	-0.054	0.177	**0.576**	0.335	-0.018	0.270	-0.332	-0.138	-0.229
LSO	0.161	0.378	0.081	0.392	0.358	-0.350	0.266	0.027	0.063
Eigenvalues	3.801	2.332	2.038	1.703	1.587	1.352	1.175	1.107	1.027
% of Variance	15.203	9.329	8.150	6.811	6.347	5.406	4.700	4.430	4.110
Cumulative %	15.203	24.533	32.683	39.494	45.841	51.247	55.947	60.377	**64.486**

Table 4.4. Results of the PCA analysis conducted on the 16 environmental variables. Values in bold represent loading scores greater than 0.50.

	$PC_{ENVIRONMENT\ 1}$	$PC_{ENVIRONMENT\ 2}$
ELE	**-0.641**	**-0.750**
RS	**-0.990**	-0.090
FL	**0.859**	0.468
FA	**0.867**	0.453
SNDPPT	-0.375	**0.901**
SLTPPT	-0.272	**-0.812**
CLYPPT	**0.802**	-0.372
CFRVOL	**-0.955**	-0.268
BIO1	**-0.711**	**0.693**
BIO5	**-0.757**	**0.648**
BIO7	**-0.858**	**0.502**
BIO10	**-0.691**	**0.710**
BIO12	**0.964**	0.255
BIO13	**0.961**	0.267
BIO15	**0.955**	-0.050
BIO16	**0.965**	0.254
Eigenvalues	10.64	4.53
% of Variance	66.50	28.35
Cumulative %	66.50	94.85

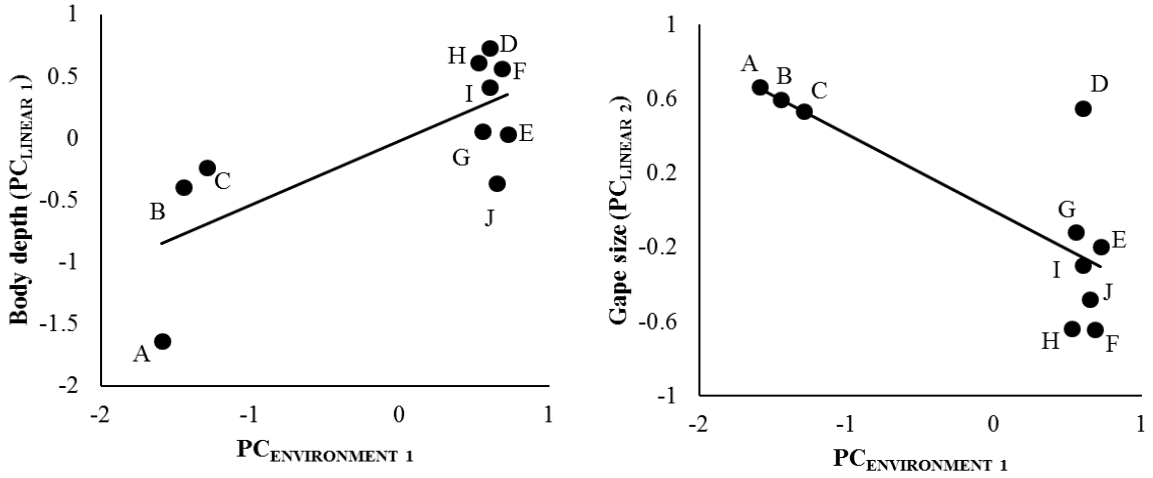


Figure 4.5. Graphs illustrating the relationship between body depth ($PC_{LINEAR\ 1}$, left) and gape size ($PC_{LINEAR\ 2}$, right) and the environmental variables ($PC_{ENVIRONMENT\ 1}$).

Multi-scale pattern analysis

Three independent MSPAs were run, namely a morphometric MSPA, environmental MSPA and a spatial and environmental MSPA in the form of a redundancy analysis (Appendix 4.2). For all three MSPAs two PC axes, explaining 51.28 %, 97.63 % and 58.89 % of the variation respectively, were retained. Environmental spatial structure was detected at broader spatial scales (MEM_{1-3}), with MEM_1 related to CLYPPT ($R^2 = 0.42$), MEM_2 related to SNDPPT, BIO1, BIO5, BIO7, BIO10 ($0.43 < R^2 < 0.77$), and MEM_3 related to CFRVOL, SLTPPT, RS, ELE, FL, FA, BIO12, BIO13, BIO15, BIO16 ($0.45 < R^2 < 0.71$) (Appendix 4.2). Spatial connectivity among samples, however, did not explain the morphological trait variation (Appendix 4.2). The MSPA redundancy analysis revealed no clear structuring when regressing the morphological traits against the environmental variables, however, most of the morphological variables appear to be associated with $MEM_{1, 3}$ (Appendix 4.3).

DISCUSSION

Multiple biotic (e.g. species composition of the invaded environment) and abiotic (e.g. environmental features) factors have been proposed to affect an organisms' invasive success (Barney & Whitlow 2008; Catford et al. 2009; Peoples et al. 2017). Here, I examined the influence environmental variables may have in shaping phenotypic variation amongst populations of an invasive species. My results revealed a significant relationship between body depth and flow regime, corroborating previous research (Langerhans 2009; Tytell et al. 2010; Lucek et al. 2014; Peoples et al. 2017). Though most previous studies measured flow in a simulated environment (i.e. lab experiments) (e.g. Langerhans 2009; Lucek et al. 2014), or collected field data (e.g. Jacquemin et al. 2013; Cerwenka et al. 2014), my study is the first to date, to use a global database with standardised measures of aquatic environmental variables (Domisch et al. 2015) to detect environmental variation. The use of standardised variables is essential when wanting to make comparisons and draw inferences between studies regarding phenotypic variation associated with invasive species in response to environmental fluctuations, as stated by Westley (2011).

Both geometric morphometrics and linear morphology supported the notion that Locality A, situated in the upper reaches of the Jan Dissels tributary, is significantly different from all other populations, however the two methods did not converge to the same morphology (geometric morphometrics: dorsal-ventral redistribution in body shape vs. linear morphology: fusiform body) (Figure 4.3, 4.4). This can likely be ascribed to the large within-population variation observed for the geometric morphometrics (Figure 4.3, 4.4). Similarly, no significant correlation was detected between any of the environmental variables and body shape ($PC_{SHAPE\ 1\text{-}4}$). The observed pattern can be explained by the fact that a fish' body shape might be related to condition rather than skeletal changes, highlighting the need for a substantially larger dataset (to compensate for the large SD) when wanting to infer body shape variation from geometric morphometric data.

My results support the prediction that comparable environments will have similar phenotypes, particularly with reference to the linear morphometric measures (Figure 4.5), with both body depth ($PC_{LINEAR\ 1}$) and gape size ($PC_{LINEAR\ 2}$) being correlated to environmental conditions. Locality A - C are characterised by high-flow environments, evident from the steeper

slope and the higher percentage of pebbles and rocks (CFRVOL) (Johnson 2004). It is thus not surprising that *M. dolomieu* from Locality A (and subsequently Locality B and C) were significantly more streamlined when compared to the other populations (Table 4.3, Figure 4.4), supporting the hypothesis of Langerhans (2008) and corroborating my second prediction that streamlined fish will be present in high-flow environments, as this body shape promotes steady swimming in fast flowing waters.

Similarly, substantial variation between localities, though not statistically significant, was observed in gape size ($PC_{LINEAR\ 2}$) (Figure 4.4), with Localities A - C, characterised by a relatively smaller gape size, clustering together. Relative gape size was found to decrease with a decrease in clay content mass (CLYPPT) (Figure 4.5), and consequently an increase in coarse fragments. Likewise, an association between gape size and CLYPPT was observed with the MSPA redundancy analysis (Appendix 4.2). Invertebrates, especially crabs, are known to prefer coarse substrate habitats situated in high flow environments (Zimmerman & Covich 2003), as present in the Jan Dissels tributary (Localities A - C). A smaller gape and consequently stronger suction (Carroll et al. 2004; Wainwright et al. 2007; Day et al. 2015; Sejdic 2016), might be advantageous in this environment as a stronger suction is required to feed on benthic prey items, such as crabs (Day et al. 2015). In contrast, a relatively large gape is better suited for ram feeding and predating on fish in the water column (Carroll et al. 2004; Wainwright et al. 2007; Day et al. 2015; Sejdic 2016). This may very well be the case in the mainstem of the Olifants River system, as numerous co-occurring fish species are present (Van der Walt et al. 2016). Moreover, the fact that *M. dolomieu* from Locality A - C have more fusiform bodies compared to the other localities further supports this notion, as steady swimming is needed to get close enough to benthic prey (Sass & Motta 2002; Higham 2011). In contrast, thrust and manoeuvrability, associated with relatively deep bodies (as seen within Localities D – J; Figure 4.4), provides *M. dolomieu* with a fast approach and attack, essential to capture prey in the water column (Sass & Motta 2002; Wainwright et al. 2007; Tran et al. 2010; Sejdic 2016).

Although previous work on invasive species has found morphological divergence to be associated with genetic variation (Lucek et al. 2014; Fitzpatrick et al. 2015), phenotypic plasticity in response to differing environmental conditions still seems to be a key component driving

morphological variation (but see Chapter 5; Langerhans et al. 2005, Langerhans 2009; Lucek et al. 2014). This might also be the case for *M. dolomieu* in the Olifants River system, as substantial genetic variation (see Chapter 2, but note that the majority of samples were not collected from the same localities used for Chapter 4), and hybridisation between *M. dolomieu* and *M. salmoides* (see Chapter 3) is present within the Olifants River system. However, a common-garden experiment would be essential to disentangle these components. Hence, future work should tease apart the genetic relationships among these localities, in order to unravel the connection between genetic variation and phenotypic plasticity. This knowledge could, in turn, help us understand how phenotypic plasticity can facilitate the successful colonisation, establishment and spread of notorious invasive species.

CHAPTER 5

GENETIC CORRELATES OF MORPHOLOGICAL VARIATION IN A CLOSED RIVER SYSTEM: INSIGHTS FROM SMALLMOUTH BASS (*MICROPTERUS DOLOMIEU*)

ABSTRACT

Phenotypic variation has often been attributed to both phenotypic plasticity and local adaptation, with various studies suggesting that these mechanisms assist in the colonisation, establishment and spread of alien invasive species in a novel invaded range Although many studies have explored the phenotypic plasticity – local adaptation continuum, few attempts have been made to do so in a natural system. Instead, conclusions are drawn from common-garden or reciprocal transplant experiments, which often cannot account for gene flow. Invasive freshwater fish provide an ideal opportunity to examine the phenotypic plasticity - local adaptation continuum and address the role of gene flow in driving phenotypic variation as they often spread into a broad range of habitats and environments, are highly mobile, express an array of morphological adaptations to their environment and exhibit extensive morphological variation. Using the invasive smallmouth bass (*Micropterus dolomieu*) in the Olifants River system, Western Cape, South Africa as model system, I predicted that there will be high levels of gene flow among localities, and that there would be no link between genetic differentiation and phenotypic variation. Seven *M. dolomieu* populations were assessed for 25 linear morphological traits before being subjected to DNA amplification for nine microsatellite loci. Population structuring, though slight, was observed among sampled localities and corresponded to three distinct sections of the river (i.e., tributary, impoundment and mainstem). As predicted, high levels of gene flow among populations were observed, but this did not erase the strong association between genetics and morphology observed for gape size and suggests that both local adaptation and phenotypic plasticity may play a key role in the invasive success of *M. dolomieu*.

INTRODUCTION

Phenotypic plasticity, an organisms' ability to express different phenotypes in response to environmental (biotic or abiotic) factors (Agrawal 2001; Davidson et al. 2011), is thought to contribute to a non-native species' invasive success (Baker 1965; Lee 2002; Ghalambor et al. 2007; Davidson et al. 2011; Matesanz et al. 2012). Intuitively, this theory is plausible as: (a) the novel invaded range may not represent the same environment as that found in the organisms' native range, and (b) invasive species are usually represented by a few colonising individuals and presumably a low genetic diversity (Davidson et al. 2011). Baker (1965) suggested that these phenotypically plastic individuals possess a 'general-purpose genotype', which allows them to successfully colonise, spread and become established in the novel environment, while simultaneously outcompeting native species when faced with fluctuating environmental conditions (Schlichting & Levin 1986; Parker et al. 2003; Davidson et al. 2011; Matesanz et al. 2012; Zenni et al. 2014). In contrast, if high genetic variation were to be present, selection toward the local optimum could occur rapidly, and phenotypic variation should arise via local adaptation rather than phenotypic plasticity (Parker et al. 2003). However, local adaptation, an individual or populations' ability to genetically change in response to local environmental cues (Williams 1966), relies on multiple founding individuals in the new environment, high levels of outcrossing (and/ or hybridisation) and novel genotype formation through elevated levels of gene flow among populations to drive phenotypic variation (Parker et al. 2003; Garant et al. 2007; Geng et al. 2007; Matesanz et al. 2012), though these are seldom the conditions in which invasive species find themselves (Dlugosch & Parker 2008).

Numerous invasion biology studies have attributed observable phenotypic variation to both phenotypic plasticity (e.g. Parker et al. 2003; Geng et al. 2007; Davidson et al. 2011; Matesanz et al. 2012) and local adaptation (e.g. Williams et al. 1995; Langerhans 2009; Dowle et al. 2015), while others, have highlighted the interplay between the two mechanisms (Lucek et al 2014; Fitzpatrick et al. 2015). A temporal pathway thus appears to exist, with phenotypic plasticity promoting colonisation and persistence in marginal environments (Crispo 2008), ultimately leading to genetic differentiation and hence, local adaptation in the novel invaded range (Ghalambor et al. 2007; Crispo 2008; Dowle et al. 2015; Fitzpatrick et al. 2012). Although many studies have tried to unravel the phenotypic plasticity versus local adaptation continuum, few have

tried to do so in natural systems, with most drawing conclusions from common-garden or reciprocal transplant experiments (e.g. Parker et al. 2003; Matesanz et al. 2012; Lucek et al. 2014; Zenni et al. 2014). These studies, however, cannot account for gene flow when wanting to unravel the adaptive basis to genetic variation (Muir et al. 2014). On the one hand, if an invasive population finds itself in a homogeneous environment and gene flow is rife, phenotypic plasticity will not be beneficial and may be lost due to natural processes such as drift, promoting local adaptation (Crispo 2008). In contrast, high gene flow between heterogenous environments, particularly if multidirectional movement is permitted, may favour phenotypic plasticity and promote individuals to adopt a jack-of-all-trades strategy, ultimately allowing individuals to maximise fitness in 'typical' environments and rapidly adapt to novel range (Crispo 2008). On the other hand, gene flow may increase a locally adapted populations' fitness by reducing the probability of inbreeding while introducing (adaptive) genetic variation (Tallmon et al. 2004; Fitzpatrick et al. 2015). However, gene flow may also constrain local adaptation (Haldane 1948; Felsenstein 1976; Slatkin 1985; Langerhans et al. 2003; Moore et al. 2007; Lucek et al. 2014), because genomic homogenisation might limit the among-population divergence to differing environments. As a result, populations could divert away from their adaptive optimum, ultimately reducing population fitness (Garcia-Ramos & Kirkpatrick 1997; Langerhans et al. 2003; Fitzpatrick et al. 2015). Hence, in situ examinations are required to account for neutral genetic processes such as gene flow, and how this may affect local adaptation and in turn, phenotypic variation (Moore et al. 2007; Muir et al. 2014).

Invasive freshwater fish provide an ideal opportunity to test for local adaptation as they often spread into a broad range of habitats and environments that may or may not resemble their native range (Westley et al. 2013). Furthermore, most are highly mobile, express an array of morphological adaptations to their environment and exhibit extensive morphological or behavioural variation (e.g., Chapter 4; Holopainen et al. 1997; Agrawal 2001; Hollander 2008). An example thereof is smallmouth bass (*Micropterus dolomieu*) in South Africa. Initially introduced from the United States of America into South Africa in 1937 for angling purposes (de Moor & Bruton 1988), *M. dolomieu* have subsequently established and spread throughout several rivers and reservoirs across the country (van Rensburg et al. 2011; Ellender & Weyl 2014). One such river is the Olifants River system, situated within the Western Cape Province of South Africa

where *M. dolomieu* has successfully established and morphologically adapted to various environmental pressures (see Chapter 4).

Invasive species provide ample opportunity to test some of these ideas as they often spread into a broad range of habitats and environments that may or may not resemble their native range (Thomson 2007). Furthermore, they are often good dispersers (Lodge 1993), express an array of morphological adaptations to their environment and exhibit extensive morphological or behavioural variation (e.g., Chapter 4; Holopainen et al. 1997; Agrawal 2001; Hollander 2008). An example hereof is the invasive smallmouth bass (*Micropterus dolomieu*) in South Africa, which after introduction from the United States of America in 1937 (de Moor & Bruton 1988) has established itself, and has spread rapidly throughout the country (van Rensburg et al. 2011; Ellender & Weyl 2014). Specifically in the Olifants River system, situated within the Western Cape Province of South Africa, *M. dolomieu* has successfully established and morphologically adapted itself in response to environmental conditions (see Chapter 4). Smallmouth bass are mobile fish known to disperse, on average, 12.1 km yr^{-1} (Sharma et al. 2009). Considering the aforementioned, and taking into account the heterogeneity of the Olifants River landscape (Chapter 4), as well as the multidirectional movement along the river (within the constraints of a linear system), I hypothesise that high mobility, through its effects on genetic variance results in a loss of adaptive phenotypes. I predict that (1) there will be high levels of gene flow present among sampled localities, (2) there will be no link between genetic differentiation and phenotypic variation. Shedding light on the species' genetic fingerprint will assist us in unravelling the interplay between phenotypic plasticity, local adaptation and gene flow, particularly in non-native species in their invaded range (García-Berthou 2007; Moore et al. 2007).

MATERIALS AND METHODS

Sampling and morphological analyses

One hundred and thirty-nine *M. dolomieu* specimens, sampled from seven of the ten localities initially sampled for Chapter 4, were used for the current chapter. The sampled localities encompassed three localities from the Jan Dissels tributary (localities A – C), Clanwilliam Dam (locality D) two from the main stem of the Olifants River (localities H – I) and one from the Ratel

tributary (locality J, Figure 5.1), respectively. Morphological data were obtained from Chapter 4. In short, 25 linear morphological traits (Table 4.1; Figure 4.2) were measured and regressed against body size. The residual scores for the seven populations used in the current chapter were retained and subjected to a principal component analysis (PCA). The resultant nine PC axes corresponded to those obtained in Chapter 4. Hence, the first two biologically relevant principal components (i.e. $PC_{LINEAR\ 1}$ - body depth and $PC_{LINEAR\ 2}$ - gape size) (Figure 4.4), were used in Chapter 5 to test for a link between morphology and genetic variation.

DNA extraction and amplification

Total genomic DNA was extracted from *M. dolomieu* muscle tissue using the NucleoSpin Tissue extraction kit (MACHEREY-NAGEL, Separations, Cape Town, South Africa) following the manufacturers protocol. To confirm the morphological identification of the collected individuals, and ensure no hybrid individuals were present within the dataset (see Chapter 3), two partial mitochondrial (mtDNA) gene regions, namely cytochrome oxidase subunit I (COI) and cytochrome b (cytb) were sequenced. PCR reactions and cycling conditions followed those stipulated in Chapter 3. All successfully sequenced individuals were checked with NCBI BLAST (cytb and COI; http://blast.ncbi.nlm.nih.gov) and BOLD (COI only; http://www.boldsystems.org/) to ensure only *M. dolomieu* were being compared in the present study.

To characterise the genetic variation, gene flow and genetic structure among the sampled populations, nine nuclear microsatellites markers (Mdo3, Mdo4, Mdo5, Mdo7, Mdo8, Mdo9, Mdo10, Mdo11, Lma21 - Colbourne et al. 1996; Malloy et al. 2000) were genotyped (Table 3.1). Three multiplex reactions were used to amplify the microsatellites and followed the protocol stipulated in Chapter 3. Microsatellites were genotyped on an automated sequencer (ABI 3730 XL DNA Analyzer, Applied Biosystems, CAF, Stellenbosch, South Africa) before being visually inspected, aligned and scored in Geneious® 10.0.2 (Biomatters, Auckland, New Zealand). Reference individuals selected from Chapter 2 and 3 were used as positive controls to ensure accurate scoring.

Microsatellite analyses

Genetic diversity and population structure

Genepop 4.2.1 (Rousset, 2008) was used to assess the genotype frequencies for deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium, while amplification errors associated with stuttering and large allele drop-out was assessed in MICROCHECKER (Van Oosterhout et al. 2006). As most of the populations were found to not comply with the HWE assumptions, FreeNA (Chapuis & Estoup 2007) was used to check for the presence of null alleles using the EM algorithm (Dempster et al. 1977).

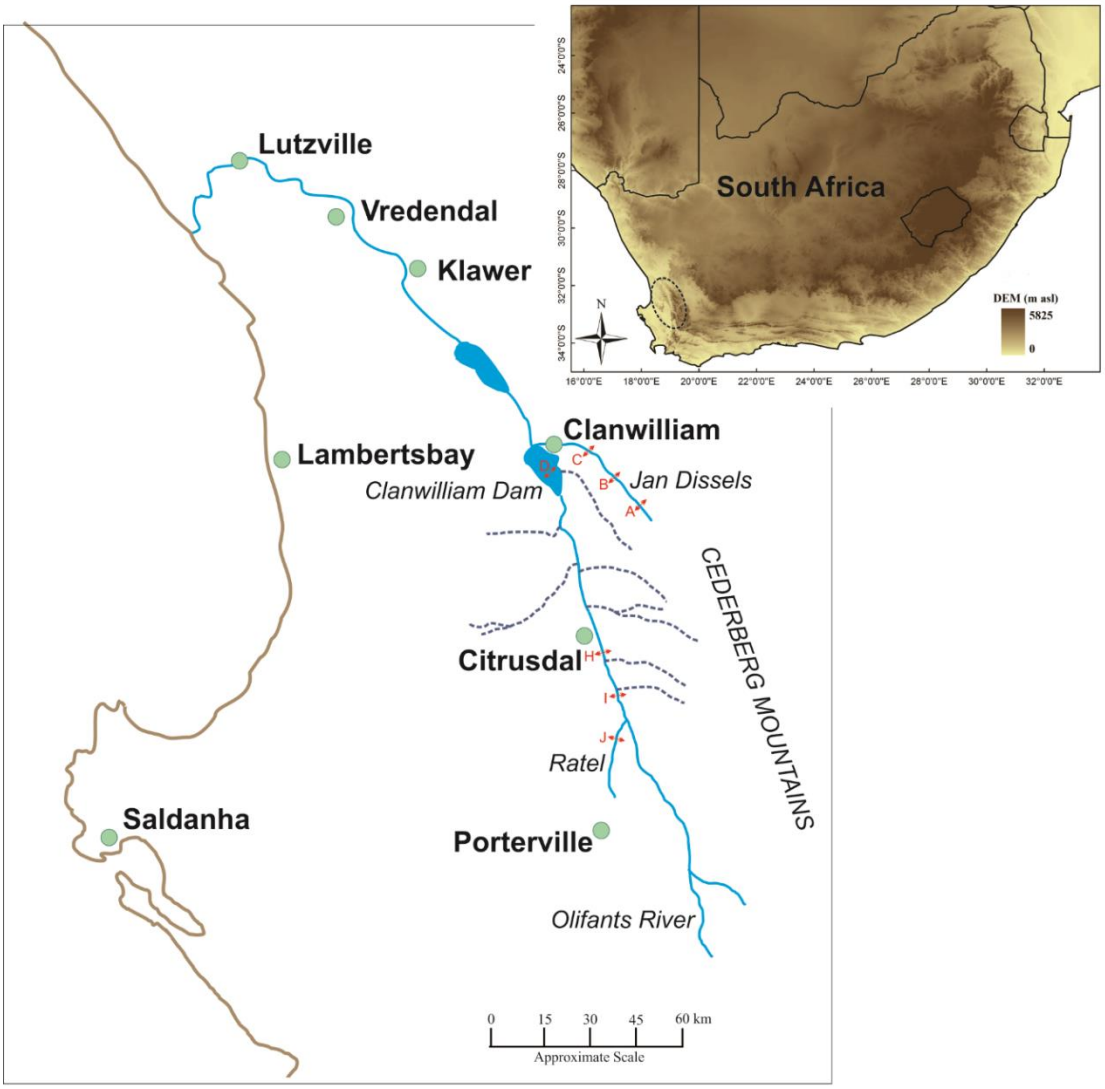


Figure 5.1. The sampling localities (A - D, H - J) within the Olifants River system of the Western Cape, South Africa. Localities A – C correspond to localities situated in the Jan Dissels tributary, Locality D represents the reservoir, while Localities H and I represent the mainstem of the Olifants River and Locality J represents the Ratel tributary.

Nuclear genomic variability within sampled populations was assessed by the allelic richness (AR), number of alleles (Na), observed heterozygosity (H_O), expected heterozygosity (H_E), and Wright's inbreeding coefficient (F_{IS}), calculated in FSTAT (Goudet 1995), Genepop (Rousset 2008) and ARLEQUIN (Excoffier & Lischer 2010). Genetic connectivity and population structure was investigated using a three-pronged approach. Firstly, Weir's (1986) F_{ST} was used to assess the genetic differentiation between sampled localities using FreeNA (Chapuis & Estoup 2007). A jackknife approach (Chapuis & Estoup 2007) implementing 1000 bootstraps, was used to assess the statistical significance. F_{ST} values range between 0 to 1, with low values signifying panmixia and higher values suggesting among-site differentiation. Genetic differentiation is considered high when values exceed 0.2, while undifferentiated populations have values near to 0 (Excoffier et al. 2005). Secondly, a Bayesian clustering approach, implemented in STRUCTURE 2.3.4 (Pritchard et al. 2000), was used to investigate the spatial population structuring along the river. Simulations were conducted using the admixture model, assuming correlated allele frequencies. Five runs, setting the number of groups (K) to range from one to eight were performed, with an initial burn-in of 50,000 followed by 250,000 MCMC iterations. The results were collated and the most probable K determined using the Evanno method (Evanno et al. 2005), as implemented Structure Harvester (Earl 2012). CLUMPP (Jakobsson & Rosenberg 2007) and Distruct (Rosenberg 2004) were used to generate and visualise the final results. Thirdly, to determine whether all specimens collected did, in fact, originate at the localities in which they were collected, and not from a neighbouring locality (i.e. a disperser), GENECLASS2 (Piry et al 2004) was used. GENECLASS2, implementing a Bayesian algorithm (Rannala & Mountain 1997) was selected to estimate the number of dispersing individuals present in the current generation. A dispersing individual is defined here as an individual not born in the population from which it was sampled, and is ultimately assigned to the most probable source population, as defined by the confidence interval (Paetkau et al. 2004). The simulation algorithm by Paetkau et al. (2004) was selected, as this method has been shown to accurately detect dispersing individuals even when all potential source populations were not sampled (Cornuet et al. 1999; Paetkau et al. 2004). The Type I error was set to 0.01, while 10 000 individuals were simulated.

Genotype-phenotype associations

To determine whether the phenotypic variation observed between the localities could be ascribed to genetic differences among localities (i.e. local adaptation), two analyses were conducted; the first being a discriminant analysis on principal components (DAPC) (Jombart et al. 2010), as implemented in the R package *adegenet* (Jombart 2008). This approach classifies individuals into clusters based on a particular trait, such as neutral genetic loci or phenotypic traits, allowing one to visualise the degree of overlap among localities at these traits (Fitzpatrick et al. 2015). Using both phenotypic dataset (i.e. the residuals) and microsatellite genotypes (all 9 microsatellites), two independent DAPC analyses were conducted on the full distribution of phenotypes and genotypes to assess the level of exchangeability of individuals among localities based on phenotypic traits and genetic similarity (Fitzpatrick et al. 2015). This was done in two parts: firstly, a PCA was conducted on all the variables (full microsatellite dataset and the morphological residuals, respectively), retaining all principal components (PCs) in order to not lose any information (as suggested by Jombart et al. 2010). Since the DAPC requires prior groupings within the dataset to be identified, the Bayesian Information Criterion (BIC) is used to identify the optimum number of clusters (*k*). Thus, the 'best' BIC value is used to determine the optimal *k* (Jombart et al. 2010). Secondly, a Discriminant Analysis is performed on the retained PCs, rendering scatterplots that clump all specimens into their most probable *k*.

Finally, to test for an association between genetic differences (microsatellite distances) and morphological variation in body depth ($PC_{LINEAR\ 1}$) and gape size ($PC_{LINEAR\ 2}$), a generalized linear model (GLM) with gaussian error distribution and identity link function was conducted using the package *lme4* (Bates et al. 2015), implemented in R. In addition to the genetic dataset (based on the pairwise F_{ST} values), geographical distance between sampled localities and environmental variables ($PC_{ENVIRONMENT\ 1}$ and $PC_{ENVIRONMENT\ 2;}$ see Chapter 4) were included in the model as explanatory variables. Firstly, to examine the multicollinearity in the dataset, the inflation factor (VIF) was calculated using the package "car" (Fox et al. 2016). Following Quinn & Keough (2002) VIF values < 10 were indicative or high collinearity. Next, multiple regression was employed to examine the influence of genetics, geographical distance and environment on morphology (i.e. $PC_{LINEAR\ 1}$ and $PC_{LINEAR\ 2}$). Backward stepwise multiple regression was used to determine the most important contributors to morphology. $PC_{LINEAR\ 1}$ and $PC_{LINEAR\ 2}$ were the dependent variables,

whereas microsatellite distances, geographical distance, $PC_{ENVIRONMENT\ 1}$ and $PC_{ENVIRONMENT\ 2}$ were the independent variables. Lastly, to identify the relative importance of each independent variable (i.e. microsatellite distances, geographical distance, $PC_{ENVIRONMENT\ 1}$, $PC_{ENVIRONMENT\ 2}$), a hierarchical partitioning analysis was computed with the package *hier.part* (Walsh & Mac Nally 2013) to further examine the explanatory variables (Quinn & Keough 2002). This analysis calculates the percentage of variance explained by each variable both jointly and independently (Chevan & Sutherland 1991; Mac Nally 1996; Walsh & Mac Nally, 2005).

RESULTS

Genetic diversity and population structure

A total of 139 specimens, collected from seven localities, were successfully sequenced (for cytb and COI) and genotyped at nine microsatellite loci. NCBI BLAST and BOLD results revealed no hybrid or introgressed individuals; hence all 139 specimens were used in subsequent analyses. With reference to the microsatellites, neither evidence of amplification errors like stuttering, large allele drop-out or null alleles were detected, nor was linkage disequilibrium observed. The majority of loci and populations did, however, not conform to Hardy-Weinberg expectations, most likely due to the significant heterozygote deficit associated with the high levels of inbreeding (F_{IS}) (Table 5.1). Observed heterozygosity (H_O) ranged from 0.150 (Locality D) to 0.950 (Locality H) across all loci, while the expected heterozygosity (H_E) ranged from 0.152 (Locality A) to 0.756 (Locality D). Marginally higher allelic richness values, based on a minimum of 15 specimens, was observed for the two mainstem localities (Locality H & I) and the Ratel tributary (Locality J) (Table 5.1). Pairwise F_{ST} values, making use of the ENA correction, revealed high levels of differentiation (> 0.2) between Locality J and all other localities, though none of these were significant ($P > 0.05$) (Table 5.2). The Bayesian clustering method implemented in STRUCTURE identified five genetic clusters ($K = 5$, LnP(D) = -2212.0; Figure 5.2), corresponding roughly to the areas sampled within the river [i.e. Jan Dissels tributary (Locality A-C), Clanwilliam Dam (Locality D), mainstem of the Olifants River (Locality H & I) and the Ratel tributary (Locality J)]. Corroborating the F_{ST} results, high levels of population differentiation was detected for Locality A and J, while high

levels of admixture were revealed for the remaining populations (Localities B-D, H, I) (Figure 5.2). Furthermore, GENECLASS2 revealed that only 75 out of 139 sampled individuals (54 % of all individuals) were correctly assigned. Thus, 46 % of *M. dolomieu* samples were found to be putative 'dispersers', belonging to a different locality than the one from which they were initially collected (Figure 5.3).

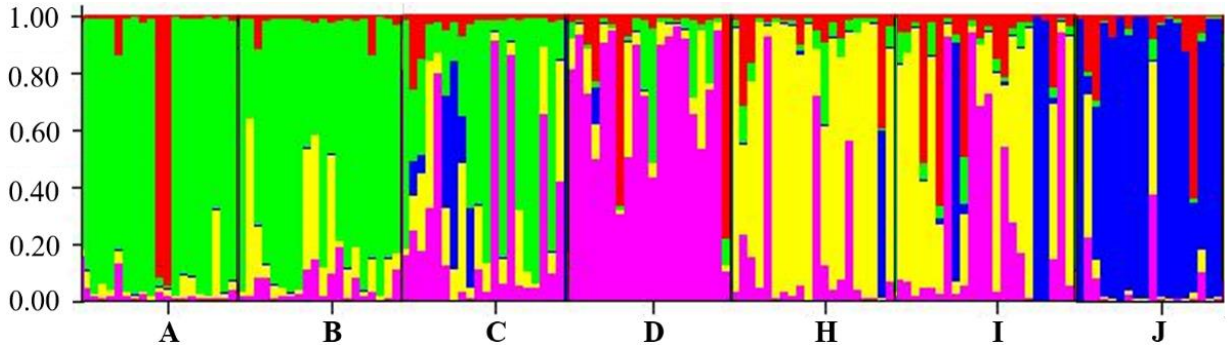


Figure 5.2. STRUCTURE plot representing all seven localities. Each line in the plot corresponds to an individual within that sampled locality. Colours represent the proportion of an individual's genotype assigned to a given genetic cluster (K = 5).

Table 5.1. Genetic diversity measures (**n** = number of genotyped individuals; **Na** = number of alleles; **AR** = allelic richness for a minimum of 15 individuals; **H_E** = expected heterozygosity, **H_O** = observed heterozygosity; **F_{IS}** = inbreeding coefficient) for nine microsatellite loci amplified for each of the seven localities. Significant values, after 1000 permutations, are represented in bold ($P < 0.05$).

		LOCALITY						
		A	**B**	**C**	**D**	**H**	**I**	**J**
Mdo3	**n**	19	20	20	20	20	22	18
	Na	4	2	3	3	3	3	2
	AR	3.659	2.000	2.978	2.700	2.700	2.999	2.000
	H_E	0.289	0.508	0.581	0.422	0.396	0.537	0.413
	H_O	0.21	0.50	0.35	0.45	0.50	0.36	0.22
	F_{IS}	**0.276**	0.016	**0.404**	-0.069	-0.271	**0.328**	0.469
Mdo4	**n**	15	20	19	17	18	21	14
	Na	3	2	3	4	2	3	4
	AR	2.998	2.000	2.997	3.620	2.000	2.999	4.000
	H_E	0.301	0.467	0.553	0.223	0.5	0.598	0.468
	H_O	0.20	0.60	0.58	0.24	0.61	0.43	0.36
	F_{IS}	0.344	-0.295	**-0.048**	-0.058	-0.23	**0.289**	0.244
Mdo5	**n**	18	19	20	20	20	21	17
	Na	3	2	3	3	3	5	4
	AR	2.733	2.000	2.995	3.000	3.000	4.561	3.824
	H_E	0.160	0.478	0.496	0.619	0.673	0.713	0.570
	H_O	0.056	0.421	0.350	0.600	0.700	0.571	0.588
	F_{IS}	**0.660**	0.122	0.300	0.032	-0.041	0.203	-0.032
Mdo7	**n**	19	20	20	20	20	21	18
	Na	5	3	4	4	4	4	4
	AR	4.673	2.999	3.978	4.000	3.978	3.991	3.778
	H_E	0.723	0.614	0.671	0.756	0.695	0.731	0.621
	H_O	0.632	0.550	0.750	0.750	0.800	0.667	0.389
	F_{IS}	0.134	0.107	-0.122	0.009	-0.156	0.089	**0.380**
Mdo8	**n**	19	20	20	20	20	22	18
	Na	2	3	4	4	5	6	3
	AR	2.000	2.700	3.400	3.694	4.377	5.729	2.954
	H_E	0.512	0.483	0.458	0.453	0.456	0.722	0.298
	H_O	0.53	0.40	0.50	0.40	0.45	0.64	0.22
	F_{IS}	-0.029	0.176	-0.095	0.119	0.014	**0.121**	**0.261**

Table 5.1 continued on next page

Table 5.1 continued

Mdo9	**n**	19	20	20	20	20	22	18
	Na	2	4	5	4	4	7	13
	AR	2.000	3.700	4.655	3.100	3.915	6.018	11.789
	H_E	0.273	0.676	0.592	0.146	0.686	0.677	0.917
	H_O	0.211	0.800	0.600	0.150	0.550	0.772	0.944
	F_{IS}	0.234	-0.190	-0.013	-0.027	**0.202**	-0.146	-0.030
Mdo10	**n**	19	20	20	20	20	22	18
	Na	3.000	2.000	2.000	3.000	2.000	2.000	4.000
	AR	2.673	2.000	2.000	2.700	1.999	2.000	3.725
	H_E	0.152	0.385	0.328	0.535	0.224	0.474	0.303
	H_O	0.105	0.400	0.300	0.650	0.150	0.545	0.333
	F_{IS}	0.314	-0.041	0.088	-0.223	0.337	-0.156	-0.103
Mdo11	**n**	18	20	20	20	19	22	18
	Na	5	2	3	5	3	3	6
	AR	4.55	2.00	3.00	4.10	3.00	3.00	5.47
	H_E	0.463	0.358	0.545	0.469	0.514	0.677	0.606
	H_O	0.50	0.45	0.65	0.55	0.63	0.55	0.61
	F_{IS}	-0.081	-0.267	-0.199	-0.177	-0.238	0.197	-0.008
Lma21	**n**	18	19	20	20	20	21	17
	Na	4	4	3	4	7	6	5
	AR	4.00	3.72	3.00	3.70	6.23	5.30	4.64
	H_E	0.705	0.603	0.504	0.483	0.7	0.612	0.41
	H_O	0.833	0.579	0.450	0.550	0.950	0.476	0.294
	F_{IS}	-0.189	0.041	0.109	-0.142	**-0.370**	0.226	0.289

Table 5.2. Pairwise F_{ST} values for each locality (A – D, H – J) based on the nine microsatellite loci. Statistical significance after 1000 bootstraps is indicated in bold ($P < 0.05$).

	A	**B**	**C**	**D**	**H**	**I**	**J**
A	-						
B	0.124	-					
C	0.112	0.056	-				
D	0.229	0.190	0.096	-			
H	0.180	0.133	0.087	0.101	-		
I	0.146	0.106	0.045	0.079	**0.030**	-	
J	0.403	0.367	0.303	0.388	0.281	0.207	-

Genotype-phenotype associations

The DAPC analyses identified seven genetic- and five morphological clusters respectively (Figure 5.4), with very little correspondence between the two. The misclassification of individuals into clusters was considerable for the genetic DAPC, with Locality J being the only locality to be represented by a cluster (Figure 5.4A). Contrastingly, individual misclassification to clusters was less pronounced for the morphological DAPC and resembled the genetic clusters obtained by the STRUCTURE analysis (Figure 5.2). However, the differentiation of Locality J was not as pronounced in the morphological DAPC when compared to the clusters retrieved by STRUCTURE (Figure 5.4B). Low VIF values were detected in the dataset (mean: 1.43, range: 1.03 – 1.87), hence multicollinearity was not considered in subsequent analyses. The backward stepwise multiple regression revealed a model with $PC_{ENVIRONMENT\ 1}$ as sole contributor to variation in $PC_{LINEAR\ 1}$ (body depth: $R^2 = 0.11$; AIC = 41.502). However, the univariate relationship between body depth and $PC_{ENVIRONMENT\ 1}$ was not significant (Table 5.3). In contrast, a model with both microsatellite distance and geographical distance was retained for $PC_{LINEAR\ 2}$ (gape size: $R^2 = 0.87$; AIC = 2.602). Both variables were highly correlated with gape size (Table 5.3). Furthermore, the hierarchical partitioning revealed that both environmental variables ($PC_{ENVIRONMENT\ 1}$ and $PC_{ENVIRONMENT\ 2}$) had the highest independent contribution with regards to body size, while microsatellite distances and geographical distance had the highest independent contribution with reference to gape size (Figure 5.5).

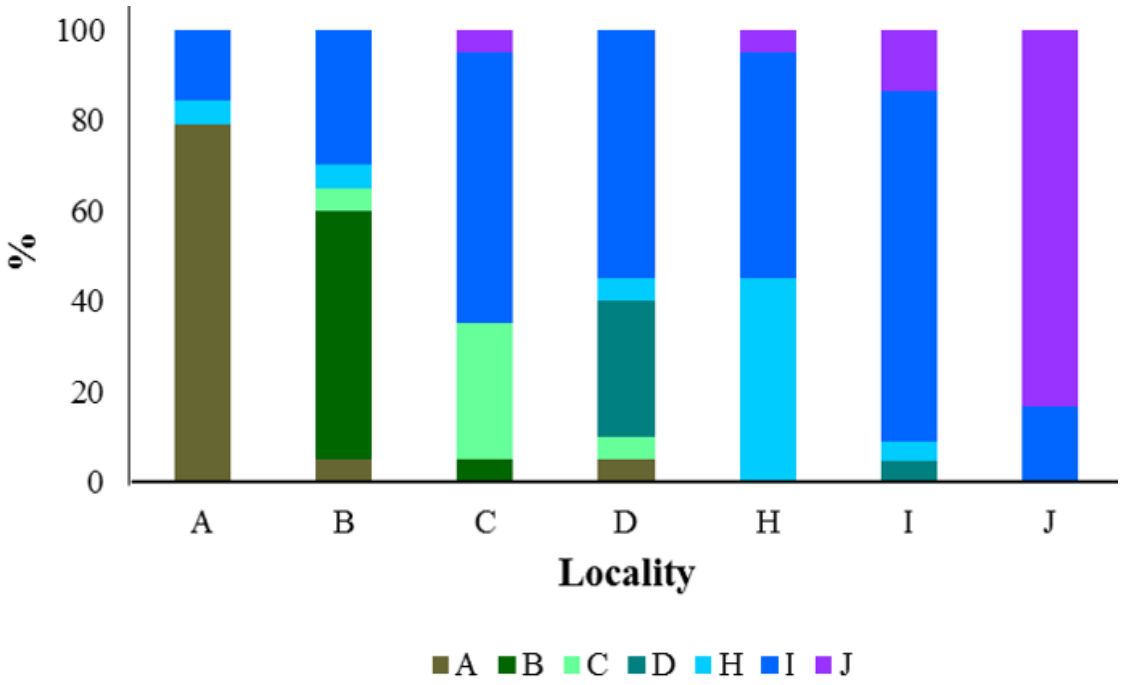


Figure 5.3. A stacked column chart representing the percentage (%) of individuals collected from and correctly assigned to each locality from which it was sampled. Letters correspond to each sampling locality.

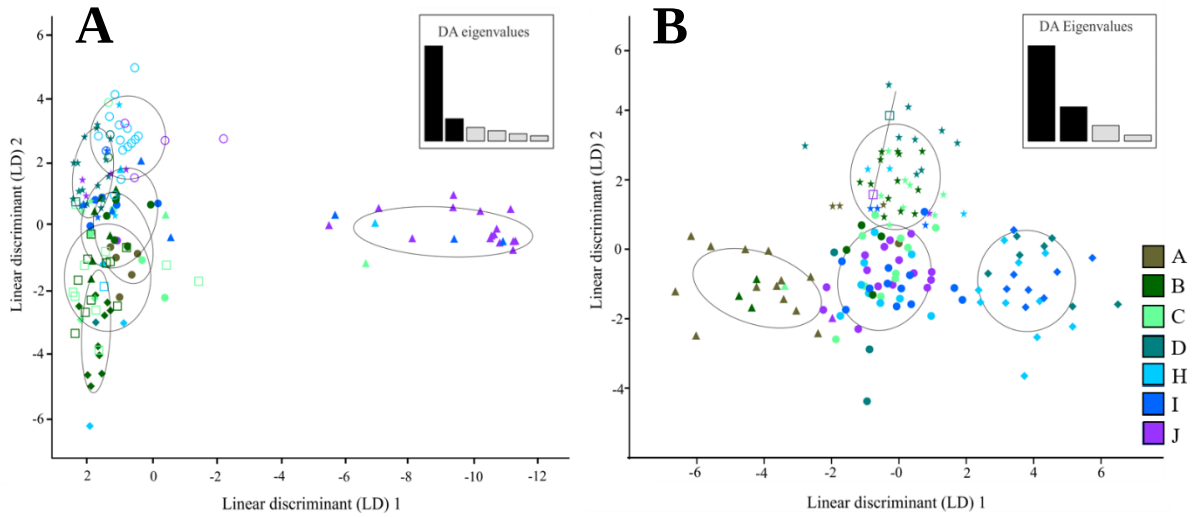


Figure 5.4. Discriminant analysis of principal components (DAPC) plots for (A) neutral genetic loci and (B) phenotypic traits. The graph represents individuals as symbols, with symbols and ellipses representing the clusters derived by the DAPC. Colours denote the sampled localities and correspond to the legend on the right. Eigenvalues for each analysis are displayed as insets.

Table 5.3. Results of the generalized linear model with gaussian error distribution and identity link function. Morphology (body depth and gape size) was used as response variables and microsatellite distances, geographical distance between sampled localities and environment ($PC_{ENVIRONMENT\ 1}$ and $PC_{ENVIRONMENT\ 2}$) as explanatory variables. Statistically significant *P*-values are indicated in bold ($P < 0.05$).

	Estimate	**SE**	***t***	***P***
		Body Depth ($PC_{LINEAR\ 1}$)		
$PC_{ENVIRONMENT\ 1}$	0.236	0.153	1.542	0.139
		Gape Size ($PC_{LINEAR\ 2}$)		
Microsatellite distances	2.693	0.494	5.454	**< 0.000**
Geographical distance	0.010	0.002	5.632	**< 0.000**

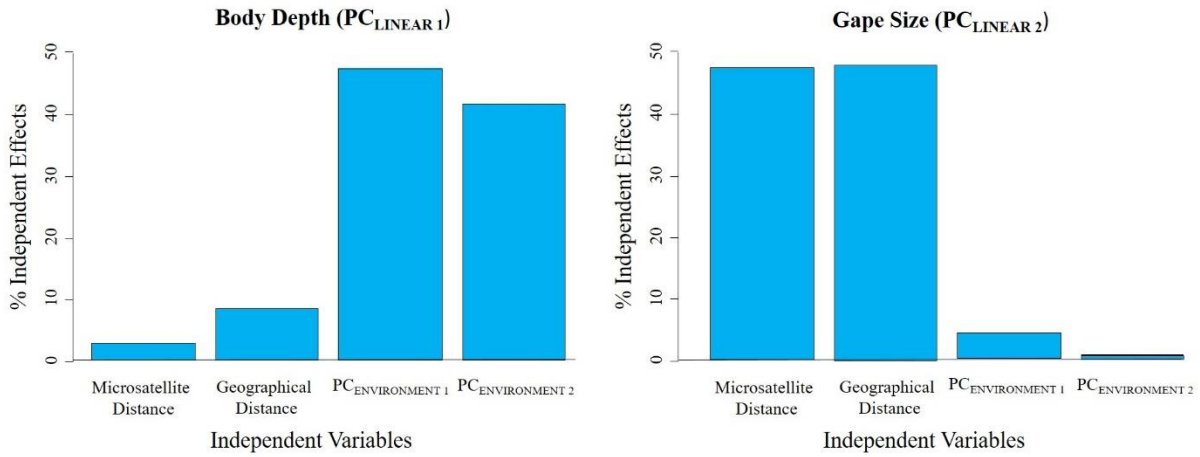


Figure 5.5. Hierarchical distribution of explanatory variables: microsatellite distances, geographical distance and environment ($PC_{ENVIRONMENT\ 1}$ and $PC_{ENVIRONMENT\ 2}$) and their percentage contribution in explaining body depth and gape size.

DISCUSSION

Rapid evolutionary change is a common phenomenon among invasive species (Brown & Marshall 1981; Thompson 1998; Mooney & Cleland 2001; Sakai et al. 2001; Lee 2002; Bossdorf et al. 2005), with many studies highlighting the role of phenotypic plasticity in facilitating the colonisation process (e.g. Yeh & Price 2004; Wund et al. 2008; Whiteley et al. 2009). Once an invasive species has colonised the novel environment, local adaptation may however take over, with gene flow introducing (adaptive) genetic variation to the newly established invasive population. In contrast, elevated levels of gene flow may also constrain local adaptation through genomic homogenisation (Haldane 1948; Felsenstein 1976; Slatkin 1985; Langerhans et al. 2003; Moore et al. 2007; Lucek et al. 2014). Hence, understanding the interplay among phenotypic plasticity, local adaptation and gene flow in invasive species is essential when wanting to unravel the species in questions' invasive success (Holopainen et al. 1997; Agrawal 2001; Moore et al. 2007; Hollander 2008; Fitzpatrick et al. 2015).

Using *M. dolomieu* as model organism to understand the phenotypic plasticity – local adaptation continuum, I predicted that (1) high levels of gene flow will be observed among sampled localities and that this would, in turn, (2) erase any genotype-phenotype association. My results corroborated the first prediction, as the STRUCTURE analysis revealed some structuring within the populations and retrieved four distinct clusters, each corresponding to the different sections of the river (i.e. Jan Dissels tributary, Clanwilliam Dam, the Olifants River mainstem and the Ratel tributary). Substantial mixing was, however, observed in Localities C, H and I (Figure 5.2). This can be attributed to the fact that individuals from the impoundment, mainstem and Ratel tributary (localities D - J) may disperse up and down the Olifants River mainstem, and into the Jan Dissels tributary when the dam overflows in winter, but as there is no fishway over the 43 m high Clanwilliam Dam wall, dispersal is unidirectional (i.e no fish from the Jan Dissels tributary can return to the impoundment). Similarly, the GENECLASS2 analysis revealed that 79% of Locality A individuals (15 out of 19) did, in fact, originate from Locality A, and were not migrants, while 83% of Locality J individuals (15 out of 18) originated from Locality J (Figure 5.3). However, as only 54% of all individuals across localities were correctly assigned to their sampled populations, it can be concluded that substantial gene flow does occur within the Olifants River system,

particularly within the mainstem, with the low levels of differentiation (i.e. F_{ST} values), too, corroborating this finding.

With reference to the genotype-phenotype association, two independent DAPC analyses were conducted. The first, a genetic DAPC, yielded seven genetic clusters, with Locality J being the only locality to represent a distinct genetic cluster (Figure 5.4A). In contrast, the second (morphological) DAPC rendered only five clusters, none of which corresponded to the genetic clusters, with Locality A being the only locality to represent a morphological cluster (Figure 5.4B). All other localities, for both the genetic- and morphological DAPC, were represented by multiple clusters (Figure 5.4). However, neither the genetic DAPC clusters, nor the morphological DAPC clusters corresponded to those obtained by STRUCTURE or the morphometric PCA analyses (conducted in Chapter 4), respectively. This may be due to the DAPC reducing the individual variation (genetic or morphological) to inter-individual or inter-population distances, thereby overlooking the within-group variation and substantially decreasing the amount of information present in the data (Jombart et al. 2010; Dufresne et al. 2014). Thus, STRUCTURE may be more powerful in detecting recently diverged population clusters as it uses the full suite of genotypic information. Furthermore, the GLM and subsequent hierarchical partitioning, revealed a significant association between genetic variation, geographical distance and gape size (Table 5.3, Figure 5.5). Although this correlation was not supported by any other analysis, this result is in accordance with previous research and hints at a genetic basis to gape morphology (Hori, 1993; Adams & Huntingford 2002; Alexander & Adams 2004; May-McNally et al. 2014). For example, Adams & Huntingford (2002) examined Arctic charr (*Salvelinus alpinus*) and observed heritable differences in gape morphology that were linked to dietary preference, while Lucek et al. (2014) attributed head shape and trophic morphology of threespine stickleback (*Gasterosteus aculeatus*) to genetic differentiation. Hence, these differences often involve not only the morphological traits associated with the detection of prey, but also encompass traits associated with capturing and/ or handling prey items, collectively referred to as trophic polymorphism (Skulason & Smith 1995). Intuitively this may be expected within my study system, as substantial variation between localities was observed in gape size, with Localities A – C displaying a relatively smaller gape than the rest of the localities (Chapter 4; Figures 4.4, 4.5). However, this may just be due to the contrast in the genetic and morphological variation between the Ratel tributary (Locality J) and the localities

within the Jan Dissels tributary (Localities A - C) and the impoundment (Locality D). The lack of association between body depth and genetic variation (Table 5.3, Figure 5.5), could be attributed to the fact that body depth is not skeletally constrained, but may rather be reliant on body condition and thus, feeding opportunities. Moreover, the influence of environmental factors on body depth (see Chapter 4) cannot be excluded.

To conclude, while some studies reveal phenotypic plasticity to be the key component driving morphological variation in invasive species (Geng et al. 2006), others attribute this variation to local adaptation (Ng et al. 2016). Yet, as highlighted by Lucek et al. (2014), both processes may have a part to play along the invasion timeline. Here, I demonstrate that invasive species can not only utilise phenotypic plasticity to colonise and persist in fluctuating environmental conditions and high levels of gene flow, but local adaptation may also be employed to facilitate in the establishment and spread of alien invasive species in a novel invaded range. An experimental approach (reciprocal transplant and/ or common-garden experiment) should be employed to further investigate the genotype-phenotype association observed in this study.

CHAPTER 6

GENERAL DISCUSSION

Extensive research has been conducted on invasive species in the hopes of establishing guidelines to facilitate our understanding of the invasion pathway and the features (organismal or environmental) that may assist in a species' invasion success (Kolar & Lodge 2001). However, despite the progress made over the past 50 years, predictions and generalisations regarding a species' invasive success remain challenging and limited (García- Berthou 2007; Leprieur et al. 2009). In this thesis, I examined the genetic and phenotypic variation within the smallmouth bass, *Micropterus dolomieu*, and the influence extrinsic factors may have on a species while adapting to the novel environment, in an attempt to add to these studies and to highlight potential pitfalls that often hinder our inferences.

Numerous studies have compared the molecular signature of invasive species in their native and invasive ranges (Guinard et al. 2003; Kolbe et al. 2004; Rollins et al. 2009; Naccarato et al. 2015), in an attempt to unravel the demographic history of the invasive populations (Ficetola et al. 2008; Gillis et al. 2009; Neilson & Stepien 2011; Gray et al. 2014). However, most studies to date have been restricted to contemporary specimens, thereby relying heavily on the premise that the historic population structure within the native range has been maintained over time and no allele frequency shifts have occurred. By including historical native specimens, I contradicted this theory in Chapter 2 and revealed substantial variation in genetic diversity between historic and contemporary native specimens, even within overlapping localities. Furthermore, not only could I unravel the introduction history of the species by including the historic native specimens, I could also examine the extensive levels of genetic differentiation within and among invasive South African *M. dolomieu* populations. Hence, I argue that historic DNA should routinely be incorporated into comparative genetic studies as it allows one to monitor the temporal changes in genetic diversity across generations (Guinard et al. 2003; Sefc et al. 2007), while increasing the chance of detecting subtle changes frequently overlooked by studies focussing only on contemporary data (Lozier & Cameron 2009). As the majority of native-invasive comparative genetic studies do not currently consider historic DNA, one cannot exclude the possibility that the temporal genetic variation accumulated over time, as observed in Chapter 2, may be a key driver

of a species' invasive success and should thus be included in subsequent genetic comparative studies.

Although I tried to minimise the error in the obtained results as far as possible, numerous limitations may have influenced my results. Firstly, sampling was the most limiting factor during this study due to the difficulties associated with collecting a sufficient sample size for comparable purposes across loci, river systems and continents. Based on historic information regarding the native *M. dolomieu* distribution range, and the hatchery and museum source populations, extensive sampling was to be conducted within these river systems to have overlapping native contemporary and historic sampling localities for Chapter 2. However, as anglers were employed to collect *M. dolomieu* in the native range, large waterbodies capable of supporting angling boats and floats were often favoured over creeks, mountain streams and rivers, limiting the amount of contemporary and historic sampling locality overlap. Furthermore, logistical issues along with financial- and time constraints hindered sample collection both within the native USA and invasive SA range.

The second limitation concerns the use of historical samples. As previously mentioned, historical DNA should be considered an invaluable tool to comparative molecular biology studies. However, these samples are not always readily available, as was the case for the historical invaded SA range. Due to the lack of samples dating back to the time of introduction, in addition to a paucity of South African museum specimens in general, the historic invaded range could not be included as it was first intended. In addition, despite having received historic native museum specimens from some musea, several (others) were hesitant to send (additional) material because of the formalin-fixed preservation method and their preconceived idea of amplification failure during DNA amplification. Hence, only 65 of the 210 specimens requested were obtained, and of these 53 were successfully amplified and scored.

Introgressive hybridisation (IH), particularly between native and introduced species, has been investigated in an array of taxa. However, very few have examined IH between two invasive species in a novel invaded range, and how introgression may facilitate the invasive success of an alien species in the novel range. By using two notorious freshwater invaders, *M. dolomieu* and *M. salmoides* as model organisms, I detected unidirectional mitochondrial introgression in Chapter 3,

despite large uncorrected pairwise mtDNA distances being observed between the two species. Similarly, simulation and empirical analyses revealed admixed individuals, extending beyond the F1 generation, suggesting viable hybrid individuals. My findings thus suggest that IH may provide a species with sufficient genetic variation to adapt to the selective pressures at hand upon introduction and supports the idea that IH may play a pivotal role in the successful establishment and spread of alien invasive species.

Chapter 3 was, however, not without its weaknesses as numerous problems were encountered during the collection of *M. salmoides*. Firstly, to test for introgressive hybridisation I had to collect *M. salmoides* and *M. dolomieu* from the same waterbodies. This proved to be quite challenging, because even though both species are abundant across the country and are often stocked together in the same artificial waterbody (i.e. reservoir), it is not common to find them co-occurring within the same natural system (i.e. river). However, as the purpose of Chapter 3 was to test for the presence of introgressive hybridisation between the two invasive species rather than the abundance thereof, opportunistic sampling across the three natural systems known to support both species (Berg, Breede and Olifants Rivers) was conducted. Fish behaviour is closely linked to diel and seasonal patterns and as such, influenced by an array of environmental factors like water temperature, water level and drought (Reebs 2002). These factors are often associated with the surface area/ volume ratio, which in turn, influences biotic interactions (e.g. predation, competition) and the chemical composition of the water (Magoulick and Kobza 2003). As an extensive drought was experienced across the Western Cape of South Africa (encompassing most of sampled river systems including the Berg, Breede and Olifants Rivers) towards the end of 2016, this too may have affected the sampling success of Chapter 3.

Organisms often face a suite of selective agents and, consequently, the observed phenotype represents a fusion of morphological traits best suited to the environment (Slatkin 1987; Robinson & Wilson 1994; Langerhans et al. 2007). Phenotypic plasticity, a mechanism thought to shield small introductory populations from strong selection by rapidly remodelling and adapting the phenotype to the new optimum, is thought to be pivotal to the successful establishment and spread of invasive species in the novel environment. Due to their restricted nature, invasive riverine fish species are ideal study organisms to examine contemporary adaptive evolution and how this may contribute to biological invasions. Using the smallmouth bass as model organism, I tested the

hypothesis stating that environmental variation drives morphological changes in phenotype in Chapter 4. My results revealed that morphology does indeed conform to environmental cues, with streamlined fish inhabit high-flow environments with percentage coarse fragments in the river bed being the best predictor of body depth. Similarly, gape size was shown to vary among localities, with fish displaying a relatively smaller gape size being present in environments with a decreased clay content mass and an increase in coarse fragments, corresponding to the prey composition associated with each substrate type. These results support the idea that similar environments will have convergent phenotypes and highlight the importance phenotypic plasticity may play in facilitating the successful colonisation, establishment and spread of invasive species. However, as large within-population morphological variation was observed, the study could have been improved by increasing the number of samples per locality. This, in turn, could have strengthened the geometric morphometric signal as well, potentially providing additional support for the hypothesis.

Phenotypic variation is often attributed to both phenotypic plasticity and local adaptation, with various studies suggesting that these mechanisms assist in the colonisation, establishment and spread of alien invasive species in a novel invaded range. Although many studies have explored the phenotypic plasticity – local adaptation continuum, few attempts have been made to do so in a natural system. Local adaptation relies on multiple founding individuals in the new environment, high levels of outcrossing (and/ or hybridisation) and novel genotype formation through elevated levels of gene flow among populations to drive phenotypic variation (Parker et al. 2003; Garant et al. 2007; Geng et al. 2007; Matesanz et al. 2012), though these are seldom the conditions in which invasive species find themselves (Dlugosch & Parker 2008). By combining morphological and genetic measures, Chapter 5 revealed *M. dolomieu* populations (separated by a mere 10 km) to display slight population structuring that corresponded to the three distinct sections of the river (i.e., tributary, impoundment and mainstem). As predicted, high levels of gene flow among populations were observed, but this did not erase the strong association between genetics and morphology observed for gape size and suggests that both local adaptation and phenotypic plasticity may play a key role in the invasive success of *M. dolomieu*.

While neutral genetic loci (mitochondrial DNA and microsatellite markers) have successfully been used in other studies to assess population structuring, gene flow and local

adaptation, next generation sequencing (NGS) may have provided a clearer, more robust picture. Furthermore, NGS has been suggested to have the potential to unravel the genetic basis behind invasiveness as it rapidly generates large amounts of sequence data and examines loci under selection (Ekblom and Galindo 2011). Although NGS was initially proposed for Chapter 5, the large within-population variation observed in the morphology of these populations (Chapter 4), led us to conclude that pooling of individuals or populations (as is often the practice with NGS) would erase any potential signal we hoped to observe between localities. Moreover, amplifying every individual for all populations was unfortunately not financially feasible. However, future research should consider employing a Genotype-by-Sequencing (GBS) approach to uncover potential loci under selection, as it is financially feasible and may significantly contribute to our understanding of mechanisms driving a species invasive success.

In summary, though several mechanisms including propagule pressure, phenotypic plasticity and elevated levels of genetic diversity are generally accepted as essential in a species' invasive success, few studies have examined these mechanisms in concert. The results of my thesis suggest that these mechanisms often facilitate or mask each other. Hence, a holistic approach should be taken when wanting to unravel an organisms' invasive success.

REFERENCES

Abell R., Thieme M.L., Revenga C., Bryer M., Kottelat M., Bogutskaya N., Coad B., Mandrak N., Balderas S.C., Bussing W. and Stiassny M.L. 2008. Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *BioScience*, *58*: 403-414.

Abernethy K. 1994. The establishment of a hybrid zone between red and sika deer (genus *Cervus*). *Molecular Ecology*, *3:* 551-562.

Aboim M.A., Mavárez J., Bernatchez L. and Coelho M.M. 2010. Introgressive hybridization between two Iberian endemic cyprinid fish: a comparison between two independent hybrid zones. *Journal of Evolutionary Biology*, *23*: 817-828.

Adams C.E. and Huntingford F.A. 2002. Inherited differences in head allometry in polymorphic Arctic charr from Loch Rannoch, Scotland. *Journal of Fish Biology*, *60*: 515-520.

Agrawal A.A. 2001. Phenotypic plasticity in the interactions and evolution of species. *Science*, *294*: 321-326.

Alexander G.D. and Adams C.E. 2004. Exposure to a common environment erodes inherited between-population trophic morphology differences in Arctic charr. *Journal of Fish Biology*, *64*: 253-257.

Allendorf F.W. 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, *5*: 181-190.

Allendorf F.W. and Leary R.F. 1988. Conservation and distribution of genetic variation in a polytypic species, the cutthroat trout. *Conservation Biology*, *2*: 170-184.

Allendorf F.W. and Lundquist L.L. 2003. Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, *17*: 24-30.

Allendorf F.W., England P.R., Luikart G., Ritchie P.A. and Ryman N. 2008. Genetic effects of harvest on wild animal populations. *Trends in Ecology & Evolution*, *23*: 327-337.

Alvarez A.C., Peterson D., Taylor A.T., Tringali M.D. and Barthel B.L. 2015. Distribution and amount of hybridization between Shoal Bass and the invasive Spotted Bass in the lower Flint River, Georgia. In *Black bass diversity: multidisciplinary science for conservation. American Fisheries Society, Symposium, 82*: 503-521.

Anderson E. and Hubricht L. 1938. Hybridization in Tradescantia. III. The evidence for introgressive hybridization. *American Journal of Botany, 25:* 396-402.

Anderson E.C. and Thompson E.A. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics, 160*: 1217-1229.

Aprahamian M.W., Smith K.M., McGinnity P., McKelvey S. and Taylor J. 2003. Restocking of salmonids - opportunities and limitations. *Fisheries Research, 62*: 211-227.

Arismendi I., Penaluna B.E., Dunham J.B., De Leaniz C.G., Soto D., Fleming I.A., Gomez-Uchida D., Gajardo G., Vargas P.V. and León-Muñoz J. 2014. Differential invasion success of salmonids in southern Chile: patterns and hypotheses. *Reviews in Fish Biology and Fisheries, 24*: 919-941.

Avise J.C., Pierce P.C., Van Den Avyle M.J., Smith M.H., Nelson W.S. and Asmussen M.A. 1997. Cytonuclear introgressive swamping and species turnover of bass after an introduction. *Journal of Heredity, 88*: 14-20.

Bagley J.C., Mayden R.L., Roe K.J., Holznagel W. and Harris P.M. 2011. Congeneric phylogeographical sampling reveals polyphyly and novel biodiversity within black basses (Centrarchidae: *Micropterus*). *Biological Journal of the Linnean Society, 104*: 346-363.

Baker H.G. 1965. Characteristics and modes of origin of weeds. In *The genetics of colonizing species*, ed. H.G. Baker & G.L. Stebbins, pp. 147-72. New York Academic Press.

Barney J.N. and Whitlow T.H. 2008. A unifying framework for biological invasions: the state factor model. *Biological Invasions, 10*: 259-272.

Barthel B.L., Cooke S.J., Svec J.H., Suski C.D., Bunt C.M., Phelan F.J.S. and Philipp D.P. 2008. Divergent life histories among smallmouth bass *Micropterus dolomieu* inhabiting a connected river - lake system. *Journal of Fish Biology, 73*: 829-852.

Bates D., Maechler M., Bolker B., Walker S., Christensen R.H.B. and Singmann H. 2015. lme4: Linear mixed-effects models using Eigen and S4, 2014. *R package version*, 1-1.

Beebee T. and Rowe G. 2005. *An Introduction to Molecular Ecology*. Oxford University Press, Oxford, UK.

Belkhir K., Borsa P., Chikhi L., Raufaste N. and Bonhomme F. 1996. GENETIX 4.02, logiciel sous Windows TM pour la génétique des populations, Laboratoire Génome, Populations, Interactions; CNRS UMR 5000; Université Montpellier II, Montpellier, France.

Beneteau C.L., Walter R.P., Mandrak N.E. and Heath D.D. 2012. Range expansion by invasion: genetic characterization of invasion of the greenside darted (*Etheostoma blennioides*) at the northern edge of its distribution. *Biological Invasions, 14:* 191-201.

Bernatchez L. and Danzmann R.G. 1993. Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of brook charr (*Salvelinus fontinalis* Mitchill). *Molecular Biology and Evolution, 10*: 1002-1002.

Berthouly-Salazar C., Hui C., Blackburn T.M., Gaboriaud C., Rensburg B.J., Vuuren B.J. and Roux J.J. 2013. Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. *Molecular Ecology*, *22*: 5793-5804.

Blackburn T.M., Pyšek P., Bacher S., Carlton J.T., Duncan R.P., Jarošík V., Wilson J.R. and Richardson D.M. 2011. A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, *26*: 333-339.

Blanchet S. 2012. The use of molecular tools in invasion biology: an emphasis on freshwater ecosystems. *Fisheries Management and Ecology*, *19*: 120-132.

Bossdorf O., Auge H., Lafuma L., Rogers W.E., Siemann E. and Prati D. 2005. Phenotypic and genetic differentiation between native and introduced plant populations. *Oecologia*, *144*: 1-11.

Bouchard F.A., Lewis S.L., Marcus C.B., McBride G.M. and Wayne M.L. 2011. Using *Drosophila melanogaster* to test the effect of multiple introductions on the ability of a non-native population to adapt to novel environments. *Evolutionary Ecology Research*, *13*: 637-646.

Bouzat J.L. 2000. The importance of control populations for the identification and management of genetic diversity. *Genetica*, *110*: 109-115.

Bowers G.M. 1905. Report of the commissioner for the year ending June 30, 1903, U.S. Commission of Fish and Fisheries. Government Printing Office, Washington, D.C.

Brown A.H.D. and Marshall D.R. 1981. Evolutionary changes accompanying colonization in plants. In *Second International Congress of systematic and evolutionary biology, the University of British Columbia, Vancouver, Canada*.

Bryant D. and Moulton V. 2004. Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, *21*: 255-265.

Buchan J.C., Archie E.A., Van Horn R.C., Moss C.J. and Alberts S.C. 2005. Locus effects and sources of error in non-invasive genotyping. *Molecular Ecology Resources*, *5*: 680-683.

Burns J.G., Di Nardo P. and Rodd F.H. 2009. The role of predation in variation in body shape in guppies *Poecilia reticulata*: a comparison of field and common garden phenotypes. *Journal of Fish Biology*, *75*: 1144-1157.

Catford J.A., Jansson R. and Nilsson C. 2009. Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity and Distributions*, *15*: 22-40.

Carroll A.M., Wainwright P.C., Huskey S.H., Collar D.C. and Turingan R.G. 2004. Morphology predicts suction feeding performance in centrarchids. *The Journal of Experimental Biology*, *207*: 3873-3881.

Cerwenka A.F., Alibert P., Brandner J., Geist J. and Schliewen U.K. 2014. Phenotypic differentiation of Ponto-Caspian gobies during a contemporary invasion of the upper Danube River. *Hydrobiologia*, *721*: 269-284.

Chapuis M.P. and Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution, 24*: 621-631.

Chevan A. and Sutherland M. 1991. Hierarchical partitioning. *The American Statistician*, *45*: 90-96.

Clement M., Posada D. and Crandall K. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, *9*: 1657-1660.

Colautti R.I. 2005. Are characteristics of introduced salmonid fishes biased by propagule pressure? *Canadian Journal of Fisheries and Aquatic Sciences*, *62*: 950-959.

Colbourne J.K., Neff B.D., Wright J.M. and Gross M.R. 1996. DNA fingerprinting of bluegill sunfish (*Lepomis macrochirus*) using (GT) n microsatellites and its potential for assessment of mating success. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*: 342-349.

Collyer M.L., Stockwell C.A., Adams D.C. and Reiser M.H. 2007. Phenotypic plasticity and contemporary evolution in introduced populations: evidence from translocated populations of white sands pupfish (*Cyrpinodon tularosa*). *Ecological Research*, *22*: 902-910.

Cooper A.M., Miller L.M. and Kapuscinski A.R. 2010. Conservation of population structure and genetic diversity under captive breeding of remnant coaster brook trout (*Salvelinus fontinalis*) populations. *Conservation Genetics*, *11*: 1087-1093.

Cornuet J.M., Piry S., Luikart G., Estoup A. and Solignac M. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, *153*: 1989-2000.

Cornuet J.M., Ravigne V. and Estoup A. 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, *11*: 401.

Cornuet J.M., Pudlo P., Veyssier J., Dehne-Garcia A., Gautier M., Leblois R., Marin J.M. and Estoup A. 2014. DIYABC v2. 0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, *30*: 1187-1189.

Crispo E. 2008. Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *Journal of Evolutionary Biology*, *21*: 1460-1469.

Darling J.A., Tsai Y.H.E., Blakeslee A.M. and Roman J. 2014. Are genes faster than crabs? Mitochondrial introgression exceeds larval dispersal during population expansion of the invasive crab *Carcinus maenas*. *Royal Society Open Science*, *1*: 140202.

Davidson A.M., Jennions M. and Nicotra A.B. 2011. Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters*, *14*: 419-431.

Day F. 1884. On races and hybrids among the Salmonidae. In *Proceedings of the Zoological Society of London*, *52*: 17-40. Blackwell Publishing Ltd.

Day S.W., Higham T.E., Holzman R. and Van Wassenbergh S. 2015. Morphology, kinematics, and dynamics: the mechanics of suction feeding in fishes. *Integrative and Comparative Biology*, *55*: 21-35.

De Moor I.J. and Bruton M.N. 1988. Atlas of alien and translocated indigenous aquatic animals in southern Africa. South African National Scientific Programmes Report 144. Foundation for Research Development and Council for Scientific and Industrial Research, Pretoria, South Africa.

Dempster A.P., Laird N.M. and Rubin D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)*, 1-38.

Dlugosch K.M. 2006. Adaptation and colonization in *Hypericum canariense*: past and present invasions of an island endemic. PhD Thesis, University of California, Santa Cruz.

Dlugosch K.M. and Parker I.M. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*: 431-449.

Dlugosch K.M., Anderson S.R., Braasch J., Cang F.A. and Gillette H.D. 2015. The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, *24*: 2095-2111.

Domisch S., Amatulli G. and Jetz W. 2015. Near-global freshwater-specific environmental variables for biodiversity analyses in 1 km resolution. *Scientific data*, *2*.

Do Prado F.D., Hashimoto D.T., Senhorini J.A., Foresti F. and Porto-Foresti F. 2012. Detection of hybrids and genetic introgression in wild stocks of two catfish species (Siluriformes: Pimelodidae): The impact of hatcheries in Brazil. *Fisheries Research*, *125*: 300-305.

Dormontt E.E., Gardner M.G., Breed M.F., Rodger J.G., Prentis P.J. and Lowe A.J. 2014. Genetic bottlenecks in time and space: reconstructing invasions from contemporary and historical collections. *PloS One*, *9*: 106874.

Dowle E.J., Morgan-Richards M., Brescia F. and Trewick S.A. 2015. Correlation between shell phenotype and local environment suggests a role for natural selection in the evolution of *Placostylus* snails. *Molecular Ecology*, *24*: 4205-4221.

Drake J.M. 2006. Heterosis, the catapult effect and establishment success of a colonizing bird. *Biology Letters*, *2*: 304-307.

Dudgeon D., Arthington A.H., Gessner M.O., Kawabata Z.I., Knowler D.J., Lévêque C., Naiman R.J., Prieur-Richard A.H., Soto D., Stiassny M.L. and Sullivan C.A. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, *81*: 163-182.

Dufresne F., Stift M., Vergilino R. and Mable B.K. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology*, *23*: 40-69.

Eales J., Thorpe R.S. and Malhotra A. 2010. Colonization history and genetic diversity: adaptive potential in early stage invasions. *Molecular Ecology*, *19*: 2858-2869.

Earl D.A. and vonHoldt B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*: 359-361.

Echelle A.A. and Echelle A.F. 1997. Genetic introgression of endemic taxa by non-natives: A case study with leon springs pupfish and sheepshead minnow. *Conservation Biology*, *11*: 153-161.

Edmands S. 2002. Does parental divergence predict reproductive compatibility? *Trends in Ecology & Evolution*, *17*: 520-527.

Ehrenfeld J.G. 2010. Ecosystem consequences of biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, *41*: 59-80.

Ekblom R. and Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, *107*: 1-15.

Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics, 5*: 435-445.

Ellender B.R., Weyl O.L. and Swartz E.R. 2011. Invasion of a headwater stream by non-native fishes in the Swartkops River system, South Africa. *African Zoology*, *46*: 39-46.

Ellender B.R. and Weyl O.L. 2014. A review of current knowledge, risk and ecological impacts associated with non-native freshwater fish introductions in South Africa. *Aquatic Invasions, 9*: 117-132.

Ellender B.R., Wasserman R.J., Chakona A., Skelton P.H. and Weyl O.L.F. 2017. A review of the biology and status of Cape Fold Ecoregion freshwater fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems*. DOI: 10.1002/aqc.2730.

Ellis E.C. and Ramankutty N. 2008. Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, *6*: 439-447.

Elton C.S. 1958. The ecology of invasions by animals and plants. The University of Chicago Press, Chicago, USA.

Estoup A., Jarne P. and Cornuet J.M. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, *11*: 1591-1604.

Estoup A. and Guillemaud T. 2010. Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, *19*: 4113-4130.

Evanno G., Regnaut S. and Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*: 2611-2620.

Excoffier L., Laval G. and Schneider S. 2005. ARLEQUIN version 3.1: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, *1*: 47-50.

Excoffier L. and Lischer H.E. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*: 564-567.

Felsenstein J. 1976. The theoretical population genetics of variable selection and migration. *Annual Review of Genetics*, *10*: 253-280.

Ferrari M.C., Ranåker L., Weinersmith K.L., Young M.J., Sih A. and Conrad J.L. 2014. Effects of turbidity and an invasive waterweed on predation by introduced largemouth bass. *Environmental Biology of Fishes*, *97*: 79-90.

Ficetola G.F., Bonin A. and Miaud C. 2008. Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology, 17*: 773-782.

Firmat C., Schliewen U.K., Losseau M. and Alibert P. 2012. Body shape differentiation at global and local geographic scales in the invasive cichlid *Oreochromis mossambicus*. *Biological Journal of the Linnean Society, 105*: 369-381.

Firmat C., Alibert P., Losseau M., Baroiller J.F. and Schliewen U.K. 2013. Successive invasion-mediated interspecific hybridizations and population structure in the endangered cichlid *Oreochromis mossambicus*. *PloS One, 8:* e63880.

Fitzpatrick S.W., Gerberich J.C., Kronenberger J.A., Angeloni L.M. and Funk W.C. 2015. Locally adapted traits maintained in the face of high gene flow. *Ecology Letters*, *18*: 37-47.

Fu Y.X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*: 915-925.

Funk W.C., Garcia T.S., Cortina G.A. and Hill R.H. 2011. Population genetics of introduced bullfrogs, *Rana* (*Lithobates*) *catesbeianus*, in the Willamette Valley, Oregon, USA. *Biological Invasions, 13*: 651-658.

Gaither M.R., Bowen B.W. and Toonen R.J. 2013. Population structure in the native range predicts the spread of introduced marine species. *Proceedings of the Royal Society B: Biological Sciences, 280*: 20130409.

Garant D.A.N.Y., Forde S.E. and Hendry A.P. 2007. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology, 21*: 434-443.

García-Berthou E. 2007. The characteristics of invasive fishes: what has been learned so far? *Journal of Fish Biology, 71*: 33-55.

Garcia-Ramos G. and Kirkpatrick M. 1997. Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, 21-28.

Garza J.C. and Williamson E.G. 2001. Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology, 10*: 305-318.

Geng Y.P., Pan X.Y., Xu C.Y., Zhang W.J., Li B. and Chen J.K. 2006. Phenotypic plasticity of invasive *Alternanthera philoxeroides* in relation to different water availability, compared to its native congener. *Acta Oecologica, 30*: 380-385.

Geng Y.P., Pan X.Y., Xu C.Y., Zhang W.J., Li B., Chen J.K., Lu B.R. and Song Z.P. 2007. Phenotypic plasticity rather than locally adapted ecotypes allows the invasive alligator weed to colonize a wide range of habitats. *Biological Invasions, 9*: 245-256.

Ghalambor C.K., McKay J.K., Carroll S.P. and Reznick D.N. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptations to new environments. *Functional Ecology, 21*: 394-407.

Gillis N.K., Walters L.J., Fernandes F.C. and Hoffman E.A. 2009. Higher genetic diversity in introduced than in native populations of the mussel *Mytella charruana*: evidence of population admixture at introduction sites. *Diversity and Distributions, 15*: 784-795.

Global Invasive Species Database: 100 of the world's worst invasive alien species, viewed 02 June 2017 from www.iucngisd.org.

Goudet J. 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity, 86*: 485-486.

Grant W.A.S. and Bowen B.W. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, *89*: 415-426.

Gray M.M., Wegmann D., Haasl R.J., White M.A., Gabriel S.I., Searle J.B., Cuthbert R.J., Ryan P.G. and Payseur B.A. 2014. Demographic history of a recent invasion of house mice on the isolated Island of Gough. *Molecular Ecology, 23*: 1923-1939.

Guinand B., Scribner K.T. and Page K.S. 2003. Genetic variation over space and time: analyses of extinct and remnant lake trout populations in the Upper Great Lakes. *Proceedings of the Royal Society of London B: Biological Sciences, 270*: 425-433.

Haldane J.B. 1948. The theory of a cline. *Journal of Genetics*, *48*: 277-284.

Hansen M.M. 2002. Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: An approach using microsatellite DNA analysis of historical and contemporary samples. *Molecular Ecology*, *11*: 1003-1015.

Hargrove J.S., Weyl O.L., Allen M.S. and Deacon N.R. 2015. Using tournament angler data to rapidly assess the invasion status of alien sport fishes (*Micropterus* spp.) in Southern Africa. *PloS One*, *10*: 0130056.

Hargrove J.S., Weyl O.L. and Austin J.D. 2017. Reconstructing the introduction history of an invasive fish predator in South Africa. *Biological Invasions*. DOI: 10.1007/s10530-017-1437-x.

Harrison A.C. 1963. The Olifants/ Doorn River system and its fishes. *Piscator,* 57: 25-39.

Harrison R.G. 1990. Hybrid zones: windows on evolutionary process. *Oxford Surveys in Evolutionary Biology*, *7*: 69-128.

Harrison R.G. and Larson E.L. 2014. Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, *105*: 795-809.

Heger T. and Trepl L. 2003. Predicting biological invasions. *Biological Invasions 5*: 313-321.

Henriques R., von der Heyden S. and Matthee C.A. 2016. When homoplasy mimics hybridization: a case study of Cape hakes (*Merluccius capensis* and *M. paradoxus*). *PeerJ, 4*: 1827.

Hermansen J.S., Haas F., Trier C.N., Bailey R.I., Nederbragt A.J., Marzal A. and Sætre G.P. 2014. Hybrid speciation through sorting of parental incompatibilities in Italian sparrows. *Molecular Ecology, 23*: 5831-5842.

Higham T.E. 2011. Feeding mechanics. *Encyclopedia of Fish Physiology: From Genome to Environment, 1*: 597-602.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. and Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology, 25*: 1965-1978.

Hijmans R.J. 2016. Introduction to the 'raster' package (version 2.5-8).

Hollander J. 2008. Testing the grain-size model for the evolution of phenotypic plasticity. *Evolution, 62*: 1381-1389.

Holopainen I.J., Tonn W.M. and Paszkowski C.A. 1997. Tales of two fish: the dichotomous biology of crucian carp (*Carassius carassius* (L.)) in northern Europe. *Annales Zoologici Fennici,* 1-22.

Hori M. 1993. Frequency dependant natural selection in the handedness of scale eating cichlid fish. *Science, 260*: 216-219.

Hubbs C.L. and Bailey R.M. 1940. A Revision of the black basses (*Micropterus* and *Huro*): with descriptions of four new forms. In *Miscellaneous Publications Museum of Zoology University of Michigan, 48*: 1-51.

Hubbs C.L. 1955. Hybridization between fish species in nature. *Systematic Zoology, 4*: 1-20.

Hui C. and Richardson D.M. 2017. Invasion Dynamics, Oxford University Press, Oxford.

Invasive Species South Africa; *Micropterus floridanus*, viewed 02 June 2017 from http://www.invasives.org.za/news-previews/item/748-florida-bass-micropterus-floridanus.

Jackson M.C., Jones T., Milligan M., Sheath D., Taylor J., Ellis A., England J. and Grey J. 2014. Niche differentiation among invasive crayfish and their impacts on ecosystem structure and functioning. *Freshwater Biology, 59*: 1123-1135.

Jackson M.C., Weyl O.L.F., Altermatt F., Durance I., Friberg N., Dumbrell A.J., Piggott J.J., Tiegs S.D., Tockner K., Krug C.B. and Leadley P.W. 2016. Chapter twelve recommendations for the next generation of global freshwater biological monitoring tools. *Advances in Ecological Research, 55*: 615-636.

Jacquemin S.J., Martin E. and Pyron M. 2013. Morphology of bluntnose minnow *Pimephales notatus* (Cyprinidae) covaries with habitat in a central Indiana watershed. *The American Midland Naturalist, 169*: 137-146.

Jakobsson M. and Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23*: 1801-1806.

Johansen-Morris A.D. and Latta R.G. 2006. Fitness consequences of hybridization between ecotypes of *Avena barbata*: hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution, 60*: 1585-1595.

Johnson S.L. 2004. Factors influencing stream temperatures in small streams: substrate effects and a shading experiment. *Canadian Journal of Fisheries and Aquatic Sciences, 61*: 913-923.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*: 1403-1405.

Jombart T., Devillard S. and Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics, 11*: 94.

Jombart T. and Ahmed I. 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*: 3070-3071.

Kalinowski S.T. 2005. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes, 5*: 187-189.

Kawamura K., Yonekura R., Ozaki Y., Katano O., Taniguchi Y. and Saitoh K. 2010. The role of propagule pressure in the invasion success of bluegill sunfish, *Lepomis macrochirus*, in Japan. *Molecular Ecology, 19*: 5371-5388.

Keller L.F. and Waller D.M. 2002. Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*: 230-241.

Kelly D.W., Muirhead J.R., Heath D.D. and Macisaac H.J. 2006. Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish waters. *Molecular Ecology*, *15*: 3641-3653.

Klingenberg C.P. 2011. MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*: 353-357.

Kolar C.S. and Lodge D.M. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution*, *16*: 199-204.

Kolbe J.J., Glor R.E., Schettino L.R., Lara A.C., Larson A. and Losos J.B. 2004. Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, *431*: 177-180.

Kovach R.P., Muhlfeld C.C., Boyer M.C., Lowe W.H., Allendorf F.W. and Luikart G. 2015. Dispersal and selection mediate hybridization between a native and invasive species. *Proceedings of the Royal Society B: Biological Sciences*, *282*: 20142454.

Lacepèdé B.G.E. 1802. Histoire naturelle des poissons. [Natural history of fish.] Chez Plasson, Paris.

Lamaze F.C., Sauvage C., Marie A., Garant D. and Bernatchez L. 2012. Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, *21*: 2877-2895.

Lamer J.T., Dolan C.R., Petersen J.L., Chick J.H. and Epifanio J.M. 2010. Introgressive hybridization between bighead carp and silver carp in the Mississippi and Illinois Rivers. *North American Journal of Fisheries Management*, *30*: 1452-1461.

Langerhans R.B., Layman C.A., Langerhans A.K. and Dewitt T.J. 2003. Habitat-associated morphological divergence in two Neotropical fish species. *Biological Journal of the Linnean Society*, *80*: 689-698.

Langerhans R.B., Layman C.A. and DeWitt T.J. 2005. Male genital size reflects a trade-off between attracting mates and avoiding predators in two live-bearing fish

species. *Proceedings of the National Academy of Sciences of the United States of America, 102*: 7618-7623.

Langerhans R.B., Chapman L.J. and DeWitt T.J. 2007. Complex phenotype–environment associations revealed in an East African cyprinid. *Journal of Evolutionary Biology*, *20*: 1171-1181.

Langerhans R.B. 2008. Predictability of phenotypic differentiation across flow regimes in fishes. *Integrative and Comparative Biology*, *48*: 750-768.

Langerhans R.B. 2009. Trade-off between steady and unsteady swimming underlies predator-driven divergence in *Gambusia affinis*. *Journal of Evolutionary Biology*, *22*: 1057-1075.

Largiadèr C.R. and Scholl A. 1996. Genetic introgression between native and introduced brown trout *Salmo trutta* L. populations in the Rhone River Basin. *Molecular Ecology*, *5*: 417-426.

Lavergne S. and Molofsky J. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences*, *104*: 3883-3888.

Leberg P.L. 2002. Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology, 11*: 2445-2449.

Lee D.S., Gilbert C.R., Hocutt C.H., Jenkins R.E., McAllister D.E. and Stauffer Jr. J.R. 1980. *Atlas of North American freshwater fishes*. North Carolina State Museum of Natural History, Raleigh, North Carolina, USA.

Lee C.E. 2002. Evolutionary genetics of invasive species. *Trends in Ecology & Evolution, 17*: 386-391.

Lee P.L., Patel R.M., Conlan R.S., Wainwright S.J. and Hipkin C.R. 2004. Comparison of genetic diversities in native and alien populations of hoary mustard (*Hirschfeldia incana* [L.] Lagreze-Fossat). *International Journal of Plant Sciences, 165*: 833-843.

Lejeusne C., Saunier A., Petit N., Béguer M., Otani M., Carlton J.T., Rico C. and Green A.J. 2014. High genetic diversity and absence of founder effects in a worldwide aquatic invader. *Scientific Reports*, *4*: 5808.

Lenormand T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*: 183-189.

Leprieur F., Beauchard O., Blanchet S., Oberdorff T. and Brosse S. 2008. Fish invasions in the world's river systems: when natural processes are blurred by human activities. *PLoS Biology*, *6*: e28.

Leprieur F., Olden J.D., Lek S. and Brosse S. 2009. Contrasting patterns and mechanisms of spatial turnover for native and exotic freshwater fish in Europe. *Journal of Biogeography*, *36*: 1899-1912.

Lindholm A.K., Breden F., Alexander H.J., Chan W.K., Thakurta S.G. and Brooks R. 2005. Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulata* in Australia. *Molecular Ecology, 14*: 3671-3682.

Lippens C., Estoup A., Hima M.K., Loiseau A., Tatard C., Dalecky A., Bâ K., Kane M., Diallo M., Sow A. and Niang Y. 2017. Genetic structure and invasion history of the house mouse (*Mus musculus domesticus*) in Senegal, West Africa: a legacy of colonial and contemporary times. *Heredity*. DOI: 10.1038/hdy.2017.18.

Lischer H.E.L. and Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*: 298-299.

Lockwood J.L., Cassey P. and Blackburn T. 2005. The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, *20*: 223-228.

Lodge D.M. 1993. Biological invasions: lessons for ecology. *Trends in Ecology & Evolution*, *8*: 133-137.

Long J., Allen M., Porak W. and Suski C.D. 2015. A historical perspective of black bass management in the United States. *American Fisheries Society Symposium*, *82*: 99-122.

Loppnow G.L., Vascotto K. and Venturelli P.A. 2013. Invasive smallmouth bass (*Micropterus dolomieu*): history, impacts, and control. *Management of Biological Invasions, 4*: 191-206.

Losos J.B., Creer D.A., Glossip D., Goellner R., Hampton A., Roberts G., Haskell N., Taylor P. and Ettling J. 2000. Evolutionary implications of phenotypic plasticity in the hindlimb of the lizard *Anolis sagrei*. *Evolution, 54*: 301-305.

Lowe W.H., Muhlfeld C.C. and Allendorf F.W. 2015. Spatial sorting promotes the spread of maladaptive hybridization. *Trends in Ecology & Evolution, 30*: 456-462.

Lozier J.D. and Cameron S.A. 2009. Comparative genetic analyses of historical and contemporary collections highlight contrasting demographic histories for the bumble bees *Bombus pensylvanicus* and *B. impatiens* in Illinois. *Molecular Ecology, 18*: 1875-1886.

Lucek K., Sivasundar A. and Seehausen O. 2014. Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. *Evolution, 68*: 2619-2632.

Lynch M. 1991. The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution, 45*: 622-629.

MacNally R. 1996. Hierarchical partitioning as an interpretative tool in multivariate inference. *Australian Journal of Ecology, 21*: 224-228.

Magoulick D.D. and Kobza R.M. 2003. The role of refugia for fishes during drought: a review and synthesis. *Freshwater Biology, 48*: 1186-1198.

Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology & Evolution, 20*: 229-237.

Malloy T.P., Van Den Bussche Jr. R.A., Coughlin W.D. and Echelle A.A. 2000. Isolation and characterization of microsatellite loci in smallmouth bass, *Micropterus dolomieu* (Teleostei: Centrarchidae), and cross-specific amplification in spotted bass, *M. punctulatus*. *Molecular Ecology, 9*: 191-195.

Marr S.M., Marchetti M.P., Olden J.D., García-Berthou E., Morgan D.L., Arismendi I., Day J.A., Griffiths C.L. and Skelton P.H. 2010. Freshwater fish introductions in mediterranean-

climate regions: are there commonalities in the conservation problem? *Diversity and Distributions*, *16*: 606-619.

Marr S.M., Ellender B.R., Woodford D.J., Alexander M.E., Wasserman R.J., Ivey P., Zengeya T. and Weyl O.L. 2017. Evaluating invasion risk for freshwater fishes in South Africa. *Bothalia-African Biodiversity & Conservation*, *47*: 1-10.

Marsh G.P. 1867. *Man and nature; or, physical geography as modified by human action*. Charles Scribner & Company, New York.

Matesanz S., Horgan-Kobelski T. and Sultan S.E. 2012. Phenotypic plasticity and population differentiation in an ongoing species invasion. *PLoS One*, *7*: 44955.

May-McNally S.L., Quinn T.P., Woods P.J. and Taylor E.B. 2014. Evidence for genetic distinction among sympatric ecotypes of Arctic char (*Salvelinus alpinus*) in south-western Alaskan lakes. *Ecology of Freshwater Fish*, *24*: 562-574.

Mayr E. 1963. *Animal species and evolution*. Cambridge, Mass: Belknap Press, Harvard.

Mayr E. and Diamond J.M. 2001. *The birds of northern Melanesia: speciation, ecology & biogeography*. Oxford University Press.

Mooney H.A. and Cleland E.E. 2001. The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences*, *98*: 5446-5451.

Moore J.S., Gow J.L., Taylor E.B. and Hendry A.P. 2007. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespine stickleback system. *Evolution*, *61*: 2015-2026.

Moyle P.B. and Light T. 1996. Biological invasions of freshwater: empirical rules and assembly theory. *Biological Conservation*, *78*: 149-161.

Muir A.P., Biek R., Thomas R. and Mable B.K. 2014. Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria*). *Molecular Ecology*, *23*: 561-574.

Muirhead J.R., Gray D.K., Kelly D.W., Ellis S.M., Heath D.D. and Macisaac H.J. 2008. Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology*, *17*: 1020-1035.

Naccarato A.M., Dejarnette J.B. and Allman P. 2015. Successful establishment of a non-native species after an apparent single introduction event: investigating ND4 variability in introduced black spiny-tailed iguanas (*Ctenosaura similis*) in southwestern Florida. *Journal of Herpetology*, *49*: 230-236.

Near T.J., Kassler T.W., Koppelman J.B., Dillman C.B. and Philipp D.P. 2003. Speciation in North American black basses, *Micropterus* (Actinopterygii: Centrarchidae). *Evolution*, *57*: 1610-1621.

Near T.J., Bolnick D.I. and Wainwright P.C. 2005. Fossil calibrations and molecular divergence time estimates in centrarchid fishes (Teleostei: Centrarchidae). *Evolution, 59*: 1768-1782.

Neff B.D., Fu P. and Gross M.R. 1999. Microsatellite evolution in sunfish (Centrarchidae). *Canadian Journal of Fisheries and Aquatic Sciences*, *56*: 1198-1205.

Nei M., Maruyama T. and Chakraborty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution*, 1-10.

Neilson M.E. and Stepien C.A. 2011. Historic speciation and recent colonization of Eurasian monkey gobies (*Neogobius fluviatilis* and *N. pallasi*) revealed by DNA sequences, microsatellites, and morphology. *Diversity and Distributions, 17*: 1-15.

Ng J., Ossip-Klein A.G. and Glor R.E. 2016. Adaptive signal coloration maintained in the face of gene flow in a Hispaniolan Anolis Lizard. *BMC Evolutionary Biology*, *16*: 193.

Nielsen E.E., Hansen M.M. and Loeschcke V. 1997. Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology*, *6*: 487-492.

Nielsen E.E.G., Bach L.A. and Kotlicki P. 2006. HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Notes, 6*: 971-973.

Novak S.J. and Mack R.N. 2005. Genetic bottlenecks in alien plant species: influences of mating systems and introduction dynamics. In: *Species Invasions: Insights into Ecology, Evolution, and Biogeography*. Eds D.F. Sax, J.J. Stachowicz & S.D. Gaines. Sinauer Associates, Sunderland, Massachusetts.

Oliveira E.F., Goulart E., Breda L., Minte-Vera C.V., de Souza Paiva L.R. and Vismara M.R. 2010. Ecomorphological patterns of the fish assemblage in a tropical floodplain: effects of trophic, spatial and phylogenetic structures. *Neotropical Ichthyology, 8*: 569-586.

Olson M.H. and Young B.P. 2003. Patterns of diet and growth in co-occurring populations of largemouth bass and smallmouth bass. *Transactions of the American Fisheries Society, 132*: 1207-1213.

Olson M.H., Young B.P. and Blinkoff K.D. 2003. Mechanisms underlying habitat use of juvenile largemouth bass and smallmouth bass. *Transactions of the American Fisheries Society, 132*: 398-405.

Paetkau D., Slade R., Burden M. and Estoup A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology, 13*: 55-65.

Paireder S., Werner B., Bailer J., Werther W., Schmid E., Patzak B. and Cichna-Markl M. 2013. Comparison of protocols for DNA extraction from long-term preserved formalin fixed tissues. *Analytical Chemistry, 439*: 152-160.

Pairon M., Petitpierre B., Campbell M., Guisan A., Broennimann O., Baret P.V., Jacquemart A.L. and Besnard G. 2010. Multiple introductions boosted genetic diversity in the invasive range of black cherry (*Prunus serotina*; Rosaceae). *Annals of Botany, 105*: 881-890.

Parker I.M., Rodriguez J. and Loik M.E. 2003. An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology, 17*: 59-72.

Parsons K.J., Robinson B.W. and Hrbek T. 2003. Getting into shape: an empirical comparison of traditional truss-based morphometric methods with a newer geometric method applied to New World cichlids. *Environmental Biology of Fishes, 67*: 417-431.

Pebesma E. 2016. Map overlay and spatial aggregation in sp.

Peoples B.K., Midway S.R., DeWeber J.T. and Wagner T. 2017. Catchment-scale determinants of nonindigenous minnow richness in the eastern United States. *Ecology of Freshwater Fish*, DOI: 10.1111/eff.12331.

Petit R.J., El Mousadik A. and Pons O. 1998. Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, *12*: 844-855.

Pfennig D.W., Wund M.A., Snell-Rood E.C., Cruickshank T., Schlichting C.D. and Moczek A.P. 2010. Phenotypic plasticity's impacts on diversification and speciation. *Trends in Ecology & Evolution*, *25*: 459-467.

Pfennig K.S., Kelly A.L. and Pierce A.A. 2016. Hybridization as a facilitator of species range expansion. *Proceedings of the Royal Society B: Biological Sciences*, *283*: 20161329.

Philipp D.P., Childers W.F. and Whitt G.S. 1983. A biochemical genetic evaluation of the northern and Florida subspecies of largemouth bass. *Transactions of the American Fisheries Society*, *112*: 1-20.

Pigliucci M., Schlichting C.D., Jones C.S. and Schwenk K. 1996. Developmental reaction norms: the interactions among allometry, ontogeny and plasticity. *Plant Species Biology*, *11*: 69-85.

Pikor L.A., Enfield K.S., Cameron H. and Lam W.L. 2011. DNA extraction from paraffin embedded material for genetic and epigenetic analyses. *Journal of Visualized Experiments*, *49*: e2763-e2763.

Pipas J.C. and Bulow F.J. 1998. Hybridization between redeye bass and smallmouth bass in Tennessee streams. *Transactions of the American Fisheries Society*, *127*: 141-146.

Piry S., Alapetite A., Cornuet J.M., Paetkau D., Baudouin L. and Estoup A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, *95*: 536-539.

Pritchard J.K., Stephens M. and Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, *155*: 945-959.

R Development Core Team. 2009. R: a language and environment of statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org.

Rannala B. and Mountain J.L. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences, 94*: 9197-9201.

Rao K. and Lakshmi K. 1999. Cryptic hybridization in marine fishes: significance of narrow hybrid zones in identifying stable hybrid populations. *Journal of Natural History, 33*: 1237-1259.

Reebs S.G. 2002. Plasticity of diel and circadian activity rhythms in fishes. *Reviews in Fish Biology and Fisheries, 12*: 349-371.

Reid D.T. and Peichel C.L. 2010. Perspectives on the genetic architecture of divergence in body shape in sticklebacks. *Integrative and Comparative Biology, 50*: 1057-1066.

Reznick D.N., Shaw F.H., Rodd F.H. and Shaw R.G. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science, 275*: 1934-1937.

Rhymer J.M. and Simberloff D. 1996. Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics, 27*: 83-109.

Ricciardi A. 2007. Are modern biological invasions an unprecedented form of global change? *Conservation Biology, 21*: 329-336.

Ricciardi A. and MacIsaac H.J. 2011. Impacts of biological invasions on freshwater ecosystems. In *Fifty years of invasion ecology: The legacy of Charles Elton* (first edition), ed. D.M. Richardson, pp.211-222. Blackwell Publishing Ltd. West Sussex, UK.

Rice W.R. 1989. Analyzing tables of statistical tests. *Evolution, 43*: 223-225.

Richards C.L., Bossdorf O., Muth N.Z., Gurevitch J. and Pigliucci M. 2006. Jack of all trades, master of some? on the role of phenotypic plasticity in plant invasions. *Ecology Letters, 9*: 981-993.

Rieseberg L.H., Kim S.C., Randell R.A., Whitney K.D., Gross B.L., Lexer C. and Clay K. 2007. Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica, 129*: 149-165.

Rius M. and Darling J.A. 2014. How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, *29*: 233-242.

Robinson B.W. and Wilson D.S. 1994. Character release and displacement in fishes: a neglected literature. *The American Naturalist, 144*: 596-627.

Robinson B.W. and Dukas R. 1999. The influence of phenotypic modifications on evolution: the Baldwin effect and modern perspectives. *Oikos*, *85*: 582-589.

Rohlf F.J. 2013. tpsDig, digitize landmarks and outlines, version 2.17. Department of Ecology and Evolution, State University of New York at Stony Brook.

Rollins L.A., Woolnough A.P., Wilton A.N., Sinclair R. and Sherwin W.B. 2009. Invasive species can't cover their tracks: using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. *Molecular Ecology, 18*: 1560-1573.

Roman J. and Darling J.A. 2007. Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, *22*: 454-464.

Rosenberg N.A. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, *4*: 137-138.

Rousset F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, *8*: 103-106.

Rutherford M.C., Mucina L. and Powrie L.W. 2006. Biomes and bioregions of southern Africa. In *The vegetation of South Africa, Lesotho and Swaziland,* eds L. Mucina & M.C. Rutherford, pp. 30-51. South African National Biodiversity Institute, Pretoria.

Sakai A.K., Allendorf F.W., Holt J.S., Lodge D.M., Molofsky J., With K.A., Baughman S., Cabin R.J., Cohen J.E., Ellstrand N.C. and McCauley D.E. 2001. The population biology of invasive species. *Annual Review of Ecology and Systematics*, *32*: 305-332.

Sardell J.M. and Uy J.A.C. 2016. Hybridization following recent secondary contact results in asymmetric genotypic and phenotypic introgression between island species of Myzomela honeyeaters. *Evolution, 70*: 257-269.

Sass G.G. and Motta P.J. 2002.The effects of satiation on strike mode and prey-capture kinematics in the largemouth bass, *Micropterus salmoides*. *Environmental Biology of Fishes, 65*: 441-454.

Schlichting C.D. and Levin D.A. 1986. Phenotypic plasticity: an evolving plant character. *Biological Journal of the Linnean Society, 29*: 37-47.

Schwarz D., Matta B.M., Shakir-Botteri N.L. and McPheron B.A. 2005. Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature, 436*: 546-549.

Scribner K.T., Page K.S. and Bartron M.L. 2000. Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries, 10*: 293-323.

Scribner K.T., Page K. and Bartron M. 2001. Life history and behavioral ecology impact rates and direction of evolutionary change in fish hybrid zones: a cytonuclear perspective. *Reviews in Fish Biology and Fisheries, 10*: 293-323.

Sefc K.M., Payne R.B. and Sorenson M.D. 2003. Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. *The Auk, 120*: 982-989.

Sefc K.M., Payne R.B. and Sorenson M.D. 2007. Single base errors in PCR products from avian museum specimens and their effect on estimates of historical genetic diversity. *Conservation Genetics, 8*: 879-884.

Sejdic A. 2016. Modulation of prey-capture kinematics and suction feeding performance in smallmouth bass, *Micropterus dolomieu*. *Masters Theses & Specialist Projects.* Paper 1573.

Sharma S., Jackson D.A. and Minns C.K. 2009. Quantifying the potential effects of climate change and the invasion of smallmouth bass on native lake trout populations across Canadian lakes. *Ecography, 32*: 517-525.

Simberloff D. and Von Holle B. 1999. Positive interactions of nonindigenous species: invasional meltdown? *Biological Invasions, 1*: 21-32.

Simberloff D. 2006. Invasional meltdown 6 years later: important phenomenon, unfortunate metaphor, or both? *Ecology Letters*, *9*: 912-919.

Simberloff D. 2009. The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, *40*: 81-102.

Simon-Bouhet B., Garcia-Meunier P. and Viard F. 2006. Multiple introductions promote range expansion of the mollusc *Cyclope neritea* (Nassariidae) in France: evidence from mitochondrial sequence data. *Molecular Ecology*, *15*: 1699-1711.

Skelton P.H. 2001. *A complete guide to the freshwater fishes of southern Africa*. Struik Publishers, Cape Town.

Skulason S. and Smith T.B. 1995. Resource polymorphisms in vertebrates. *Trends in Ecology & Evolution*, *10*: 366-370.

Slatkin M. 1985. Rare alleles as indicators of gene flow. *Evolution*, *39*: 53-65.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science*, *236*: 787-793.

Sloss B.L., Jennings M.J., Franckowiak R. and Pratt D.M. 2008. Genetic identity of brook trout in Lake Superior south shore streams: potential for genetic monitoring of stocking and rehabilitation efforts. *Transactions of the American Fisheries Society*, *137*: 1244-1251.

Smith G.R. 1992. Introgression in fishes: significance for paleontology, cladistics, and evolutionary rates. *Systematic Biology*, *41*: 41-57.

Stepien C.A., Murphy D.J. and Strange R.M. 2007. Broad- to fine-scale population genetic patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great Lakes and beyond: an interplay of behaviour and geography. *Molecular Ecology*, *16*: 1605-1624.

Storz J., Sabatino S., Hoffman F., Gering E., Moriyama H., Ferrand N., Monteiro B. and Nachman M. 2007. The molecular basis of high-altitude adaptation in deer mice. *PLoS Genetics*, *3*: e45.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*: 585-595.

Tallmon D.A., Luikart G. and Waples R.S. 2004. The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution, 19*: 489-496.

Taylor G.C. and Weyl O.L.F. 2013. Validation of annulus formation in otoliths of largemouth bass *Micropterus salmoides* outside their native range. *Journal of Applied Ichthyology, 29*: 381-386.

Thomas C.D. 2015. Rapid acceleration of plant speciation during the Anthropocene. *Trends in Ecology and Evolution, 30*: 448-455.

Thompson J.N. 1998. Rapid evolution as an ecological process. *Trends in Ecology & Evolution, 13*: 329-332.

Thompson G.D., Robertson M.P., Webber B.L., Richardson D.M., Le Roux J.J. and Wilson J.R.U. 2011. Predicting the subspecific identity of invasive species using distribution models: *Acacia saligna* as an example. *Diversity and Distributions, 17*: 1001-1014.

Tran H.Q., Mehta R.S. and Wainwright P.C. 2010. Effects of ram speed on prey capture kinematics of juvenile Indo-Pacific tarpon, *Megalops cyprinoides*. *Zoology, 113*: 75-84.

Twyford A.D. and Ennos R.A. 2012. Next-generation hybridization and introgression. *Heredity, 108*: 179-189.

Tytell E.D., Borazjani I., Sotiropoulos F., Baker T.V., Anderson E.J. and Lauder G.V. 2010. Disentangling the functional roles of morphology and motion in the swimming of fish. *Integrative and Comparative Biology, 50*: 1140-1154.

Vähä J. and Primmer C. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridisation scenarios and with different numbers of loci. *Molecular Ecology, 15*: 63-72.

Van der Walt J.A, Weyl O.L.F., Woodford D.J. and Radloff F.G.T. 2016. Spatial extent and consequences of black bass (*Micropterus* spp.) invasion in a Cape Floristic Region river basin. *Aquatic conservation: Marine and Freshwater Ecosystems, 26*: 736-748.

Vander Zanden M.J., Olden J.D., Thorne J.H. and Mandrak N.E. 2004. Predicting occurrences and

impacts of smallmouth bass introductions in north temperate lakes. *Ecological Applications, 14*: 132-148.

Van Kleunen M., Weber E. and Fischer M. 2010. A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters, 13*: 235-245.

Van Oosterhout C., Weetman D. and Hutchinson W.F. 2006. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes, 6*: 255-256.

Van Rensburg B.J., Weyl O.L.F., Davies S.J., van Wilgen N.J., Spear D., Chimimba C.T. and Peacock F. 2011. Invasive vertebrates of South Africa. In *Biological invasions: Economic and environmetal costs of alien plant, animal, and microbe species*, ed. D. Pimentel, pp. 326-378. CRC Press, Boca Raton, Florida.

Vilà M., Weber E. and Antonio C.M. 2000. Conservation implications of invasion by plant hybridization. *Biological Invasions, 2*: 207-217.

Von der Heyden S. and Connell A. 2012. Evidence of hybridisation within the genus Chrysoblephus and conserved nuclear sequences of South African sparids (Teleostei: Sparidae). *African Journal of Marine Science, 34*: 505-510.

Wainwright P.C., Alfaro M.E., Bolnick D.I. and Hulsey C.D. 2005. Many-to-one mapping of form to function: a general principle in organismal design? *Integrative and Comparative Biology, 45*: 256-262.

Wainwright P., Carroll A.M., Collar D.C., Day S.W., Higham T.E. and Holzman R.A. 2007. Suction feeding mechanics, performance, and diversity in fishes. *Integrative and Comparative Biology, 47*: 96-106.

Walsh C. and MacNally R. 2013. hier. part: hierarchical partitioning. R package version 1.0-3.

Ward R.D., Zemlak T.S., Innes B.H., Last P.R. and Hebert P.D. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360*: 1847-1857.

Wares J.P., Hughes A.R. and Grosberg R.K. 2005. Mechanisms that drive evolutionary change: insights from species introductions and invasions. In *Species invasions: insights into ecology, evolution, and biogeography*, pp. 229-257. Sinauer & Associates, Sunderland, Massachusetts.

Webster M.S. and Reichart L. 2005. Use of microsatellites for parentage and kinship analyses in animals. *Methods in Enzymology*, *395*: 222-238.

Wellband K.W. and Heath D.D. 2017. Plasticity in gene transcription explains the differential performance of two invasive fish species. *Evolutionary Applications*, DOI: 10.1111/eva.12463.

West-Eberhard M.J. 2003. *Developmental plasticity and evolution*. Oxford University Press.

Westley P.A. 2011. What invasive species reveal about the rate and form of contemporary phenotypic change in nature. *The American Naturalist, 177*: 496-509.

Westley P.A., Conway C.M. and Fleming I.A. 2012. Phenotypic divergence of exotic fish populations is shaped by spatial proximity and habitat differences across an invaded landscape. *Evolutionary Ecology Research*, *14*: 147-167.

Westley P.A., Ward E.J. and Fleming I.A. 2013. Fine-scale local adaptation in an invasive freshwater fish has evolved in contemporary time. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*: 20122327.

Weyl O.L.F., Ellender B.R., Woodford D.J. and Jordaan M.S. 2013. Fish distributions in the Rondegat River, Cape Floristic Region, South Africa, and the immediate impact of rotenone treatment in an invaded reach. *African Journal of Aquatic Sciences, 38*: 201-209.

Whiteley A.R., Gende S.M., Gharrett A.J. and Tallmon D.A. 2009. Background matching and color-change plasticity in colonizing freshwater sculpin populations following rapid deglaciation. *Evolution*, *63*: 1519-1529.

Whitmore D.H. and Butler W. 1982. Interspecific hybridization of smallmouth and Guadalupe bass (*Micropterus*): Evidence based on biochemical genetic and morphological analyses. *The Southwestern Naturalist, 27*: 99-106.

Whitmore D.H. and Hellier T.R. 1988. Natural hybridization between largemouth and smallmouth bass (*Micropterus*). *Copeia, 1988*: 493-496.

Williams G.C. 1966. *Adaptation and Natural Selection*. Princeton University Press. Princeton, New York.

Wilson C.C. and Hebert P.D. 1993. Natural hybridization between Arctic char (*Salvelinus alpinus*) and lake trout (*S. namaycush*) in the Canadian Arctic. *Canadian Journal of Fisheries and Aquatic Sciences, 50*: 2652-2658.

Wilson C.C. and Bernatchez L. 1998. The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Molecular Ecology, 7*: 127-132.

Wilson J.R.U., Dormontt E.E., Prentis P.J., Lowe A.J. and Richardson D.M. 2009. Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology & Evolution, 24*: 136-144.

Winemiller K.O. 1991. Ecomorphological diversification in lowland freshwater fish assemblages from five biotic regions. *Ecological Monographs, 61*: 343-365.

Wund M.A., Baker J.A., Clancy B., Golub J.L. and Foster S.A. 2008. A test of the "flexible stem" model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *The American Naturalist, 172*: 449-462.

Yeh P.J. and Price T.D. 2004. Adaptive phenotypic plasticity and the successful colonization of a novel environment. *The American Naturalist, 164*: 531-542.

Yokogawa K. 2013. Morphological variations in bluegill, *Lepomis macrochirus*, with particular emphasis on growth-related changes. *Ichthyological Research, 60*: 48-61.

Yonekura R., Kawamura K. and Uchii K. 2007. A peculiar relationship between genetic diversity and adaptability in invasive exotic species: bluegill sunfish as a model species. *Ecological Research, 22*: 911-919.

Zenni R.D., Lamy J.B., Lamarque L.J. and Porté A.J. 2014. Adaptive evolution and phenotypic plasticity during naturalization and spread of invasive species: implications for tree invasion biology. *Biological invasions*, *16*: 635-644.

Zimmerman J.K. and Covich A.P. 2003. Distribution of juvenile crabs (*Epilobocera sinuatifrons*) in two Puerto Rican headwater streams: effects of pool morphology and past land-use legacies. *Archiv für Hydrobiologie*, *158*: 343-357.

Appendix 2.1. A detailed description of specimens obtained from various museums, including the specimen origin, collection date, specimen abbreviation corresponding to that used in Table 2.1 and the main text (Chapter 2), museum responsible for the specimen and its corresponding accession number.

Country	State	Sampled Locality	Drainage System	Collection Date	Specimen Abbrev.	Material Supplied By	Accession #	Notes
USA	Maryland	Monocacy River	Potomac River	1941	PO_1	ANSP	ANSP 95683	fry
USA	Maryland	Monocacy River	Potomac River	1941	PO_2	ANSP	ANSP 95683	fry
USA	Maryland	Monocacy River	Potomac River	1941	PO_3	ANSP	ANSP 95683	fry
USA	Maryland	Plummer Is., Maryland.	Potomac River	1930	PO_4	NMNH	USNM 284083	fin snip & bits of gillraker; my have been exposed to arsenic (As), mercury (Hg), lead (Pb)
USA	Virginia	Shenandoah River	Shenandoah River	1934	SH_1	NMNH	USNM 102132	Muscle tissue
USA	Virginia	Shenandoah River	Shenandoah River	1935	SH_2	NMNH	USNM 93780	Muscle tissue
USA	West Virginia	Shenandoah River	Shenandoah River	1936	SH_3	NMNH	USNM 100694	Muscle tissue
USA	Virginia	Shenandoah River	Shenandoah River	1933	SH_4	NMNH	USNM 104928	Muscle tissue
USA	Ohio	Mosquito Creek	Mosquito Creek	1938	MO_1	OSUM	OSUM 3568	Muscle tissue
USA	Ohio	Mosquito Creek	Mosquito Creek	1938	MO_2	OSUM	OSUM 3568	Muscle tissue
USA	Ohio	Auglaize River	Auglaize River	1940	AU_1	OSUM	OSUM 3814	Muscle tissue
USA	Ohio	Auglaize River	Auglaize River	1940	AU_2	OSUM	OSUM 3814	Muscle tissue
USA	Ohio	Auglaize River	Auglaize River	1940	AU_3	OSUM	OSUM 3942	Muscle tissue
USA	Ohio	Pusheta Creek	Auglaize River	1941	AU_4	OSUM	OSUM 4343	Muscle tissue
USA	Ohio	Pusheta Creek	Auglaize River	1941	AU_5	OSUM	OSUM 4343	Muscle tissue
USA	Ohio	Lake Erie	Lake Erie	1941	LE_1	OSUM	OSUM 4272	Muscle tissue
USA	Ohio	Lake Erie	Lake Erie	1941	LE_2	OSUM	OSUM 4272	Muscle tissue
USA	Ohio	Lake Erie	Lake Erie	1941	LE_3	OSUM	OSUM 4272	Muscle tissue
USA	Ohio	White Oak Creek	Ohio River	1930	OH_1	OSUM	OSUM 10834	Muscle tissue
USA	Ohio	White Oak Creek	Ohio River	1930	OH_2	OSUM	OSUM 10834	Muscle tissue
USA	Ohio	White Oak Creek	Ohio River	1930	OH_3	OSUM	OSUM 10834	Muscle tissue
USA	Michigan	Grosse Isle shore, Detroit river	Detroit River	1935	DE_1	UMMZ	UMMZ 243459	Muscle tissue

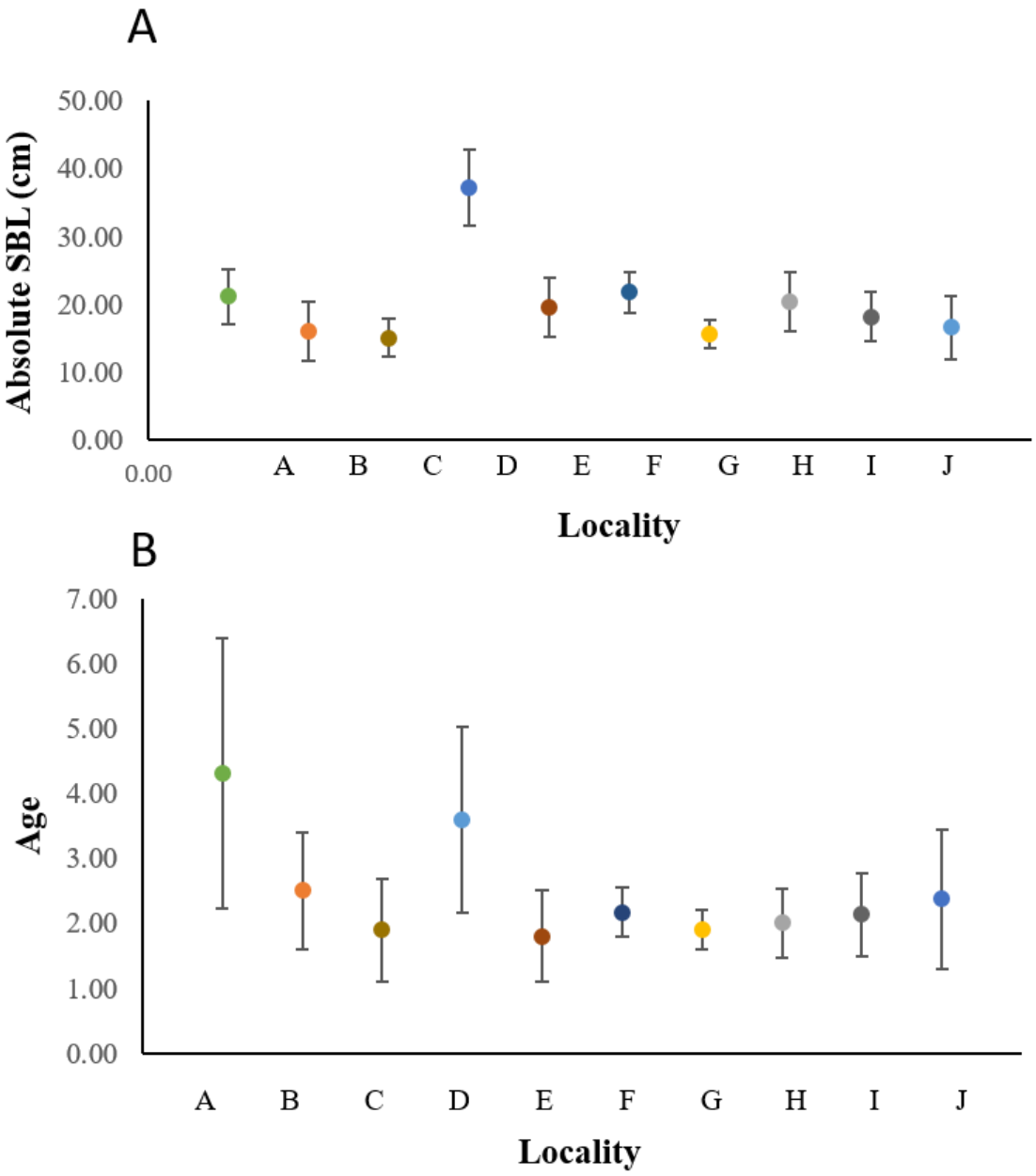
USA	Michigan	Grosse Isle shore, Detroit river	Detroit River	1935	DE_2	UMMZ	UMMZ 243459	Muscle tissue
USA	Michigan	Grosse Isle shore, Detroit river	Detroit River	1935	DE_3	UMMZ	UMMZ 243459	Muscle tissue
USA	Michigan	Grosse Isle shore, Detroit river	Detroit River	1935	DE_4	UMMZ	UMMZ 243459	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1935	DE_5	UMMZ	UMMZ 243226	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1935	DE_6	UMMZ	UMMZ 243226	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1935	DE_7	UMMZ	UMMZ 243077	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1935	DE_8	UMMZ	UMMZ 243077	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1935	DE_9	UMMZ	UMMZ 243077	Muscle tissue
Canada	Ontario	Detroit River	Detroit River	1940	DE_10	UMMZ	UMMZ 130878	Muscle tissue
Canada	Ontario	Detroit River	Detroit River	1940	DE_11	UMMZ	UMMZ 130878	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1934	DE_12	UMMZ	UMMZ 243009	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1934	DE_13	UMMZ	UMMZ 243009	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1934	DE_14	UMMZ	UMMZ 243009	Muscle tissue
USA	Michigan	Detroit River	Detroit River	1934	DE_15	UMMZ	UMMZ 243009	Muscle tissue
USA	Ontario	Detroit River	Detroit River	1940	DE_16	UMMZ	UMMZ 130896	Muscle tissue
USA	Ontario	Detroit River	Detroit River	1940	DE_17	UMMZ	UMMZ 130896	Muscle tissue
USA	Ontario	Detroit River	Detroit River	1940	DE_18	UMMZ	UMMZ 130896	Muscle tissue
USA	New York	Otselic River	Susquehanna River	1935	SU_1	UMMZ	UMMZ 109652	Muscle tissue
USA	New York	Otselic River	Susquehanna River	1935	SU_2	UMMZ	UMMZ 109652	Muscle tissue
USA	New York	Otselic River	Susquehanna River	1935	SU_3	UMMZ	UMMZ 109652	Muscle tissue
USA	New York	Susquehanna River	Susquehanna River	1935	SU_4	UMMZ	UMMZ 109759	Muscle tissue
USA	New York	Susquehanna River	Susquehanna River	1935	SU_5	UMMZ	UMMZ 109759	Muscle tissue
USA	New York	trib Rondout River to Hudson River	Hudson River	1936	HU_1	UMMZ	UMMZ 114240	Muscle tissue
USA	New York	trib Rondout River to Hudson River	Hudson River	1936	HU_2	UMMZ	UMMZ 114240	Muscle tissue
USA	New York	trib Rondout River to Hudson River	Hudson River	1936	HU_3	UMMZ	UMMZ 114240	Muscle tissue
USA	New York	trib Rondout River to Hudson River	Hudson River	1936	HU_4	UMMZ	UMMZ 114240	Muscle tissue

USA	New York	Allegheny River	Alleghany River	1937	AL_1	UMMZ	UMMZ 180878	Muscle tissue
USA	New York	Allegheny River	Alleghany River	1937	AL_2	UMMZ	UMMZ 180878	Muscle tissue
USA	New York	Allegheny River	Alleghany River	1937	AL_3	UMMZ	UMMZ 180878	Muscle tissue
USA	New York	Fall Creek, trib. to Cayuga Lake, Etna	Fall Creek	1931	FC_1	UMMZ	UMMZ 94455	Muscle tissue
USA	New York	Fall Creek, trib. to Cayuga Lake, Etna	Fall Creek	1931	FC_2	UMMZ	UMMZ 94455	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR2	SAIAB	AC09 B425	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR3	SAIAB	AC09 B955	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR4	SAIAB	AC09 B875	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR5	SAIAB	AC09 B992	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR6	SAIAB	AC09 B994	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR7	SAIAB	AC09 B977	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR8	SAIAB	AC09 B960	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR9	SAIAB	AC09 B964	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR10	SAIAB	AC09 B982	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR11	SAIAB	AC09 B978	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR12	SAIAB	AC09 B971	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR13	SAIAB	AC09 B997	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR14	SAIAB	AC09 B970	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR15	SAIAB	AC09 B984	Muscle tissue
SA	Eastern Cape	Elandsjacht Dam	Krom	2012	KR16	SAIAB	AC09 B963	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU1	SAIAB	OW14-965	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU2	SAIAB	OW14-985	Muscle tissue

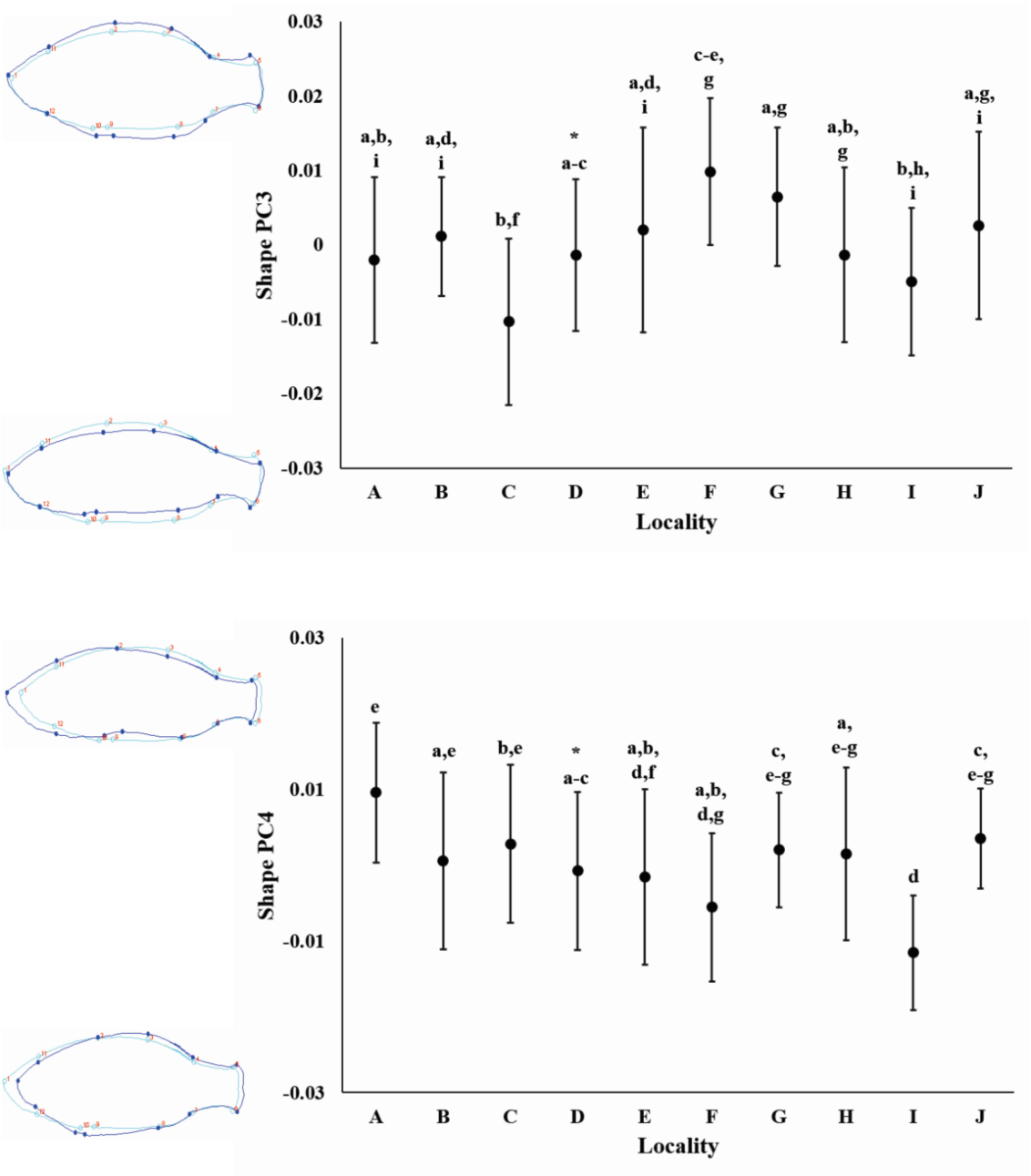
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU3	SAIAB	OW14-979	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU4	SAIAB	OW14-941	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU5	SAIAB	OW14-835	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU6	SAIAB	OW14-828	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU7	SAIAB	OW14-791	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU8	SAIAB	OW14-700	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU9	SAIAB	OW14-798	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU10	SAIAB	OW14-688	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU11	SAIAB	OW14-684	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2014	BU12	SAIAB	OW14-808	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU13	SAIAB	OW14-737	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU14	SAIAB	OW14-735	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU15	SAIAB	OW14-742	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU16	SAIAB	OW14-724	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU17	SAIAB	OW14-686	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU18	SAIAB	OW14-797	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU19	SAIAB	OW14-796	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU20	SAIAB	OW14-675	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU21	SAIAB	OW14-702	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU22	SAIAB	OW14-744	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU23	SAIAB	OW14-705	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU24	SAIAB	OW14-782	Muscle tissue

SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU25	SAIAB	OW14-732	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU26	SAIAB	OW14-746	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU27	SAIAB	OW14-756	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU28	SAIAB	OW14-738	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU29	SAIAB	OW14-733	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU30	SAIAB	OW14-739	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU31	SAIAB	OW14-799	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU32	SAIAB	OW14-715	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU33	SAIAB	OW14-704	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU34	SAIAB	OW14-762	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU35	SAIAB	OW14-727	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU36	SAIAB	OW14-690	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU37	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU38	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU39	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU40	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU41	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU42	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU43	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU44	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU45	SAIAB		Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU46	SAIAB		Muscle tissue

SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU47	SAIAB	Muscle tissue
SA	Eastern Cape	Rooikranz Dam	Buffalo River	2015	BU48	SAIAB	Muscle tissue



Appendix 4.1. Size (A) and age (B) structure for the ten *M. dolomieu* populations sampled in Chapter 4. Means and standard deviations are shown.



Appendix 4.2. Graphs illustrating the mean and standard deviation for each locality with regards to shape PC3 and shape PC4. The fish outline drawings depict the variation in fish body shape for each PC axis, with the light blue line representing the average shape for all fish, while the dark blue line represents the upper and lower body shape extremity (scale factor set to 0.08 and - 0.08, respectively). Different alphabetical letters above the bars indicate values with statistically significant differences (Bonferroni *post hoc* test; $P < 0{\cdot}05$), while identical letters indicate no significant differences among means. The source population is depicted with an asterisks (*).

Appendix 4.3. The results from the MSPA analysis conducted on morphological traits (left), environmental variables (middle) and the redundancy analysis (RDA) (right). Eigenvalues are represented by insets. MEM_{1-3} in the RDA analysis are colour coded according to their strength of association (R^2 value) with the independent environmental variables (insert bottom right of RDA figure).